Depth Extraction from 3D-Integral Images

submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Silvia MANOLACHE

DE MONTFORT UNIVERSITY, UK

October 9, 2001

Abstract

The development of full colour, fatigue free, autostereoscopic 3D imaging systems exhibiting parallax in all directions has been a constant pursuit of the scientific community, particularly in the last century. The main methods of producing 3D images, stereoscopy, holography, and integral imaging, are reviewed and the 3D integral imaging system developed at De Montfort University is described in detail. The generation of depth maps from 3D images is of major importance for modern electronic display systems to enable real and computer generated objects to be incorporated in the same frame. Depth estimation can be equally well used in applications like 3D modelling of natural objects, 3D remote handling and quality control, virtual studio and 3D telepresence.

A mathematical model of the 3D-integral imaging system, which provides the analytic expression of the imaging operator, is derived. The imaging operator is used to formulate depth extraction as an inverse problem of the image formation process. Existing regularization methods as well as a novel hierarchical adaptive constrained 3D reconstruction algorithm are employed to derive intensity maps from the planar recording of an integral image. Finally, the depth (contour) map of the object space is constructed from the intensity map using a modified elastic Durbin-Willshaw net.

Acknowledgements

I would like to express my deep gratitude and address my thanks to:

- my supervisors: Dr. Amar Aggoun, Prof. Malcolm McCormick, and Mr. Neil Davies, for their invaluable advice, help, trust, and support throughout my PhD studies.

- Prof. S.Y. Kung, for his painstaking supervision during my visit at Princeton University and excellent advice in research matters.

- the members of the 3D-Imaging Group, for their support.

- the support staff at De Montfort University, especially to Mrs. Sheila Hayto, and Mrs. Ann Clarke.

- my husband, Dr. Marcian Cirstea, for his patient love and encouragement.

- Miss Oana Spulber, Dr. and Mrs. Andrei and Anca Dinu, Prof. Shankar Madathil, Dr. Merlyne De Souza, Miss Doris Ressmann, Miss Souhila Fekkai, Mr. and Mrs. Subhas and Sahana Bose, Mr. and Mrs. Shyam and Swati Hardikar, and Mr. Paolo Fioravanti, for their constant friendship, kindness, and support.

- my parents, for their endless love and unfailing support.

Silvia Manolache,

Leicester, October 2001

To my parents,

with love

Contents

Chapter 1

Introduction

The development of 3D-imaging systems has been a constant pursuit of the scientific as well as of the entertainment community in the new technological era. Interest in this field is motivated by the numerous possible commercial and industrial applications, ranging from robotic vision for remote control, medical visualisation to leisure products. Recent technological progress and the better understanding of human visual perception have lead to the development of systems able to produce images with unprecedented levels of 'three-dimensionality'.

Imaging systems employing incoherent light sources have many advantages, not the least being application in a non-specialised environment. Integral photography (IP), which is based on using microlens arrays for capturing images, is the main nonstereoscopic method available in this area and was first proposed by Lippmann ([53]) in 1908. It creates true colour three dimensional images which do not require special viewing spectacles, display continuous parallax and do not flip. The method was further investigated and improved by Ives in order to overcome the pseudoscopy of the image produced by Lippmann's system ([40]). Burckhardt et al. ([13], [14]) proposed a new technique and reported a theoretical study which attempted to determine the optimum parameters of an IP system. Okoshi also studied the optics of the IP in the particular case of lenticular integral photography, which uses arrays of cylindrical surface microlenses. Recently, vigorous efforts have been made to develop an integral system for television applications ([19], [73]), and pick-up and display systems have been reported ([11], [73]).

The 3D integral imaging system developed at De Montfort University produces

3D images which are true optical models rather than multiple view stereoscopy, thereby demonstrate continuous parallax in all directions. This 3D imaging system, if successfully engineered, has enormous implications for leisure and industrial applications. A major potential application is 3D television. Generation of depth maps is essential if real and/or computer generated objects are to be integrated within integral 3D TV images. Depth estimation can be equally well used in applications like 3D modelling of natural objects, 3D remote handling, industrial inspection systems for quality control, virtual studio and 3D telepresence. Integral 3D images store the depth map of the scene encoded optically in the intensity profiles of the many micro-intensity distributions at the recording plane, on the photographic plate. As a result, the depth map can be generated from a single integral image, avoiding the need for the use of multiple cameras or displacement of the camera to capture multi-views of the scene.

The research project aims to derive an algorithm to reconstruct the composition of the object space and generate the depth map from integral 3D images produced with the 3D integral imaging system developed at De Montfort University.

Common approaches estimate depth information by a triangulation of corresponding image points in stereoscopic image pairs with known stereoscopic camera parameters. The present research addresses the depth extraction problem in a non-stereoscopic context and proposes a new approach: depth estimation as an inverse problem of the imaging process. The main objectives to be achieved within the project are:

- To develop a mathematical model of the integral imaging process, including the derivation of the point spread function of the optical process.
- To identify existing regularization methods applicable to depth extraction as an inverse problem of the imaging process and to obtain intensity maps of the object space using these methods.
- To develop an algorithm able to provide an accurate intensity map of the object space at a good resolution.
- To derive the depth map of the object space from the intensity map.

1.1 The scope of the thesis and original contributions

The integral imaging technique uses microlens arrays to record the three dimensional integral images onto a photosensitive detector, e.g. photographic film. Information is stored two dimensionally, and to recover the spatial (3D) object space information microlens arrays are used at replay. The 3D data is therefore contained entirely in a planar recording.

The first step on the way from raw recorded integral data to a depth map of the object space is to decipher the mechanism of spatial information encoding in the process of integral image formation and recording ([56], [57]). The mathematical study of the integral image formation and recording, performed for a comprehensive set of possible microlens configurations, proves that depth information is contained both in the number of recorded micro-intensity distributions for a single physical point and in the spread of such a micro-intensity distribution. This information is expressed by the total point spread function of the process.

The optical and recording processes are fully determined by the imaging operator associated with the point spread function, which is linear, continuous, and contains depth information. The derivation of this operator allows depth extraction to be formulated as the inverse problem of image formation. A solution of this inverse problem provides the intensity distribution in the imaged scene, which allows the recovery of the composition of the object space and the subsequent depth estimation. In this way, in the present approach, depth extraction becomes associated with object space reconstruction.

Due to the loss of information that occurs during the direct process – image formation and recording, the inverse problem is ill-posed in the continuous formulation and very ill-conditioned in the discretised version. The use of the pseudo-inverse of the imaging operator is thus inefficient, and additional constraints, coming from the physics of the problem, need to be imposed to obtain realistic solutions. Existing regularization methods are employed to produce solutions of the object space reconstruction in the form of intensity maps of the scene. The solutions thus obtained reproduce the shape and position of the objects, but suffer from not being always

positive and from lack of robustness to perturbations caused by interpolation at the sampling points. Moreover, the resolution in the object space is limited by the number of pixels in the image.

To overcome these difficulties, a new hierarchical depth extraction algorithm is proposed. An adaptive constrained 3D-reconstruction (AC3DR) algorithm, based on the use of sigmoid functions, is derived to produce positive bounded regularized solutions. To increase the precision of the reconstruction, a hierarchical multiresolution scheme is developed to detect and produce local maps of the high intensity regions of the object space using AC3DR. Finally, the depth (contour) map of the object space is constructed from the intensity map using a modified elastic Durbin-Willshaw net.

The following aspects are investigated in this research work:

1. A mathematical model of the 3D unidirectional (lenticular) integral imaging system, based on cylindrical surface microlenses, for close and remote imaging conditions has been derived. Both geometrical and physical aspects of the optical process have been analysed, emphasising the point depth relevance for the final recorded sampled data set. The point spread function of the process is obtained.

2. A mathematical model of the 3D omnidirectional integral imaging system has been developed for a few alternative configurations of the system, depending on the type of microlenses used (square based, circular based, or hexagonal based spherical surface microlenses) and on the microlens packaging (rectangular net or honeycomb). Both close and remote imaging conditions have been considered. The relationship between the recorded intensity distributions and the point depth is particularly analysed. The point spread function is derived.

3. Depth extraction is formulated as the inverse problem of the imaging process using the point spread function of the system. A computational approach of the discretisation of the problem has been described.

4. Existing regularization methods have been employed to produce physically meaningful solutions of the inverse problem. Tests have been performed on

a set of computer generated integral images and the results obtained using various methods have been assessed and compared.

5. A new adaptive constrained 3D reconstruction algorithm (AC3DR), based on the use of sigmoid functions and able to provide positive bounded solutions, has been derived. It is hierarchically applied according to a hierarchical multiresolution strategy (HMS) to obtain higher accuracy and higher resolution intensity maps. The depth map of the object space is then extracted from the intensity map using a neural network technique, namely a Durbin-Willshaw elastic net, which has been modified to suit the requirements of the problem. Tests have been performed on a comprehensive set of integral images.

The thesis encompasses the following novel developments and original contributions:

1. A novel mathematical model of the 3D unidirectional integral imaging system ([56]).

2. A novel mathematical model of the 3D omnidirectional integral imaging system ([57]).

3. A new approach to depth extraction, formulated as the inverse problem of the imaging process ([58], [59]).

4. A new adaptive constrained 3D-reconstruction (AC3DR) algorithm. It can be used as a regularization method for obtaining positive bounded solutions of any ill-posed inverse problem ([58], [59]).

5. A new hierarchical multiresolution strategy which allows stable finer resolution reconstructions at low computational expense. It can be used in conjunction with any of the existing regularization methods ([58], [59], [60])).

6. A new method of deriving a depth map from an intensity map using a weighted Durbin-Willshaw elastic net. It can be used in conjunction with any of the existing regularization methods ([59]).

1.2 Outline of the thesis

Chapter 2 contains a brief historical overview of three main streams in the 3D imaging field: stereoscopy, holography, and integral imaging. Depth extraction approaches in stereoscopy are summarised. The place and potential of integral imaging is analysed in comparison with other existing 3D imaging systems.

Chapter 3 comprises a detailed description of the integral imaging system developed at De Montfort University in terms of camera operation and optical process (image formation, transmission, and recording phases), and the mathematical model of the 3D-unidirectional (lenticular) integral camera system. The geometrical aspects of image formation analysed are: image coordinate computation, disparity, imaging fields, number of recorded intensity distributions for a single point. The point spread function for different segments of the camera and for the entire microlens system, using physical optics and mathematical analysis tools, is computed. In addition, the relationship between point spread dimensions and point depth is derived for both close and remote imaging conditions, and illustrative simulation results are presented. Depth influence on image formation and recording is extensively analysed.

Chapter 4 contains a generalisation of the mathematical model of the 3D unidirectional integral camera system to the 3D-omnidirectional integral camera system. Various camera configurations are considered, depending on the microlens type used (square based, circular based, hexagonal based microlenses) and the microlens array packaging (rectangular net or honeycomb). The geometry of the image formation and recording, as well as the point spread function for close and remote conditions are derived. The relationship between point depth and the recorded micro-intensity distributions is discussed.

In Chapter 5, depth extraction is formulated as the inverse problem of the imaging process, which proves ill-posed in the continuous form and ill-conditioned in the discretised variant. Approximate solutions satisfying additional constraints coming from the physics of the problem are searched, using so-called *regularization methods.* The Chapter comprises an extensive overview of existing regularization methods applicable to depth extraction.

In Chapter 6, the regularization methods reviewed in Chapter 5 are applied to a set of unidirectional computer generated integral images, and the reconstructions of the object space thus obtained are assessed and compared.

Chapter 7 proposes a novel algorithm for object space reconstruction and depth extraction from the planar (2D) recorded data set of a 3D-integral image. This algorithm has three components:

1. An adaptive constrained 3D-reconstruction (AC3DR) algorithm based on the use of sigmoid functions.

2. A hierarchical multiresolution strategy (HMS) which employs the adaptive constrained algorithm to obtain higher resolution and higher accuracy intensity maps of the object space.

3. A depth map derivation from the intensity map using a modified Durbin-Willshaw elastic net.

Chapter 8 formulates a list of conclusions, and possible further developments of this research work are suggested.

Chapter 2

Overview of 3D-Imaging Techniques

2.1 Introduction

Three dimensional imaging is enjoying growing interest which has produced an increasing number of commercial and industrial applications, ranging from remote control in robotic vision and medical imaging to entertainment applications. Many of the methods of producing images with associated depth perception have a long history, but the present technological level combined with a better understanding of the human vision mechanisms have lead to the development of systems with unprecedented levels of '3D reality' ([64]). The present chapter contains a brief historical overview of three main streams in the field: stereoscopy, holography, and integral imaging. It is beyond the aim of this thesis to present an exhaustive review of all existing stereoscopic and holographic systems. The place and potential of integral imaging is analysed in comparison with other existing 3D imaging systems. The manner in which depth extraction is approached in stereoscopy and various computational procedures are presented in section 2.2.3. To the best of our knowledge, to date, the only systematical study of depth extraction from integral images has been reported in [59], [58], [60] and this is comprehensively presented in Chapters 5-7.

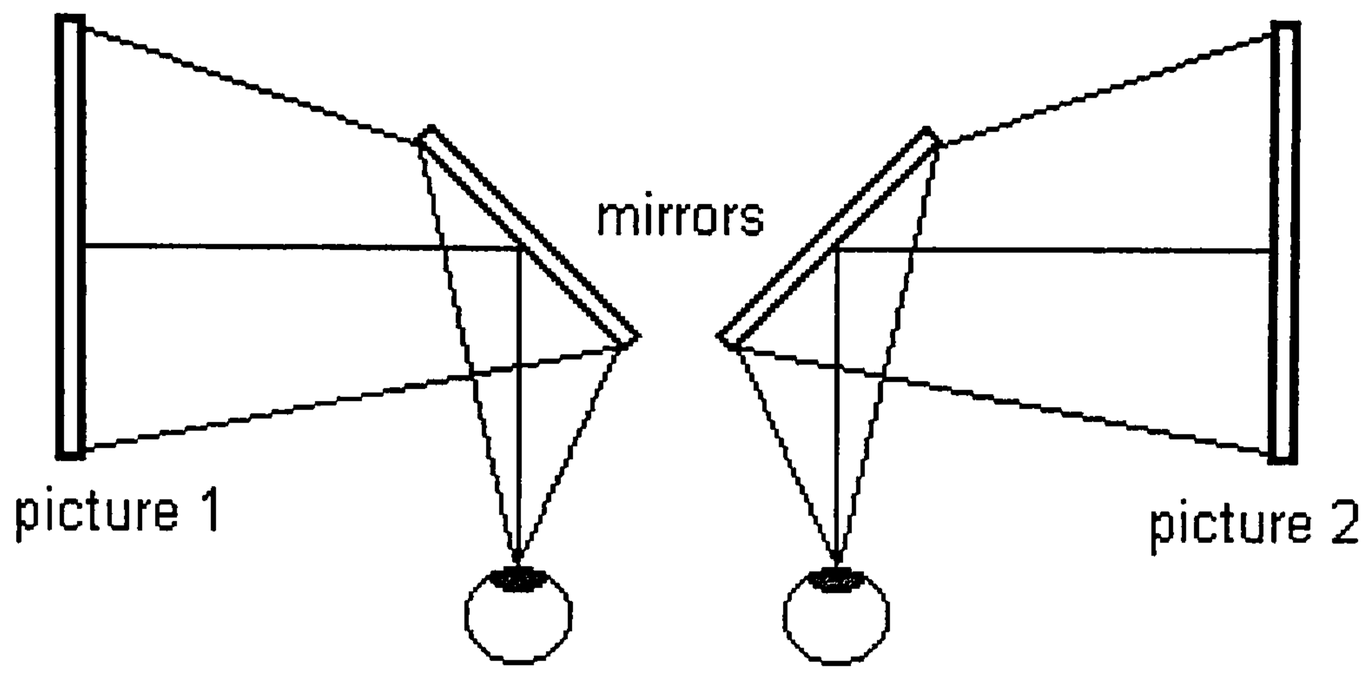


Figure 2.1: Wheatstone's stereoscope ([74]).

2.2 Stereoscopy: from Wheatstone's stereoscope to virtual studios

2.2.1 Early stereoscopy

Stereoscopic imaging is the oldest type of 3D-imaging and dates back to the middle of the XIX-th century when Wheatstone and Brewster developed the first stereoscopes ([93], [74]). The principle of the stereoscope is based on the fact that viewing an image of an object with the left eye and another image of the same object taken from a slightly changed direction with the right eye causes the brain to fuse the two images and produce a three dimensional perception. Depth sensation is thus created on the grounds of binocular parallax and convergence. The operation of a stereoscope is depicted in Figure 2.1 ([74]).

There are several methods of taking binocular stereopictures, which can be classified as follows ([74]):

1. Two pictures are taken successively by a single camera.

2. Two pictures are taken simultaneously by two gauged cameras.

3. Two pictures are taken simultaneously by a special camera. Such a camera may be a single-lens camera ([74]) able to collect by itself the binocular information using a large lens. An alternative system, called 'front-lens beam-splitter'

([74]), collects stereo information at appropriately separated positions and records two pictures in half the regular size.

The same principles can be equally well applied to TV systems, and many binocular stereoscopic television models have been proposed and tested. An extensive classification of these systems can be found in [74]. There are two main groups of systems ([74]):

1. direct-view systems

2. projection-type systems

The main difference between the two groups is that the systems of the second type do not send the image pairs directly to the viewer, but they first project them onto a specially coated screen ([74]). The methods employed to obtain the stereoscopic effect are essentially the same:

1. Two-colour method ([74]). A colour cathode ray tube displays or projects on a special screen the right-eye and the left-eye images using two different colours, e.g. the left one in red, and the right one in green. The viewer observes the image through a pair of glasses having a red lens on the left eye and a green lens on the right eye. The colours of the glasses act in this system like a direction selective device.

2. Polarisation method ([74]). Two cathode ray tubes display or project on a special screen the right-eye and the left-eye images, polarised orthogonally to each other through polaroid filters and then synthesised by a half mirror, as shown in Figure 2.2. The viewers attain a depth sensation when looking at the image through a pair of polaroid glasses.

3. Time division method ([74]). The two images are sent alternately to the two eyes. A cathode ray tube displays or projects on a screen alternately the right-eye and the left-eye images, while the viewer wears a pair of shutter mechanisms in front of his eyes. This mechanism shields each eye alternately, presenting the left-eye image to the left eye and the right-eye image to the right eye.

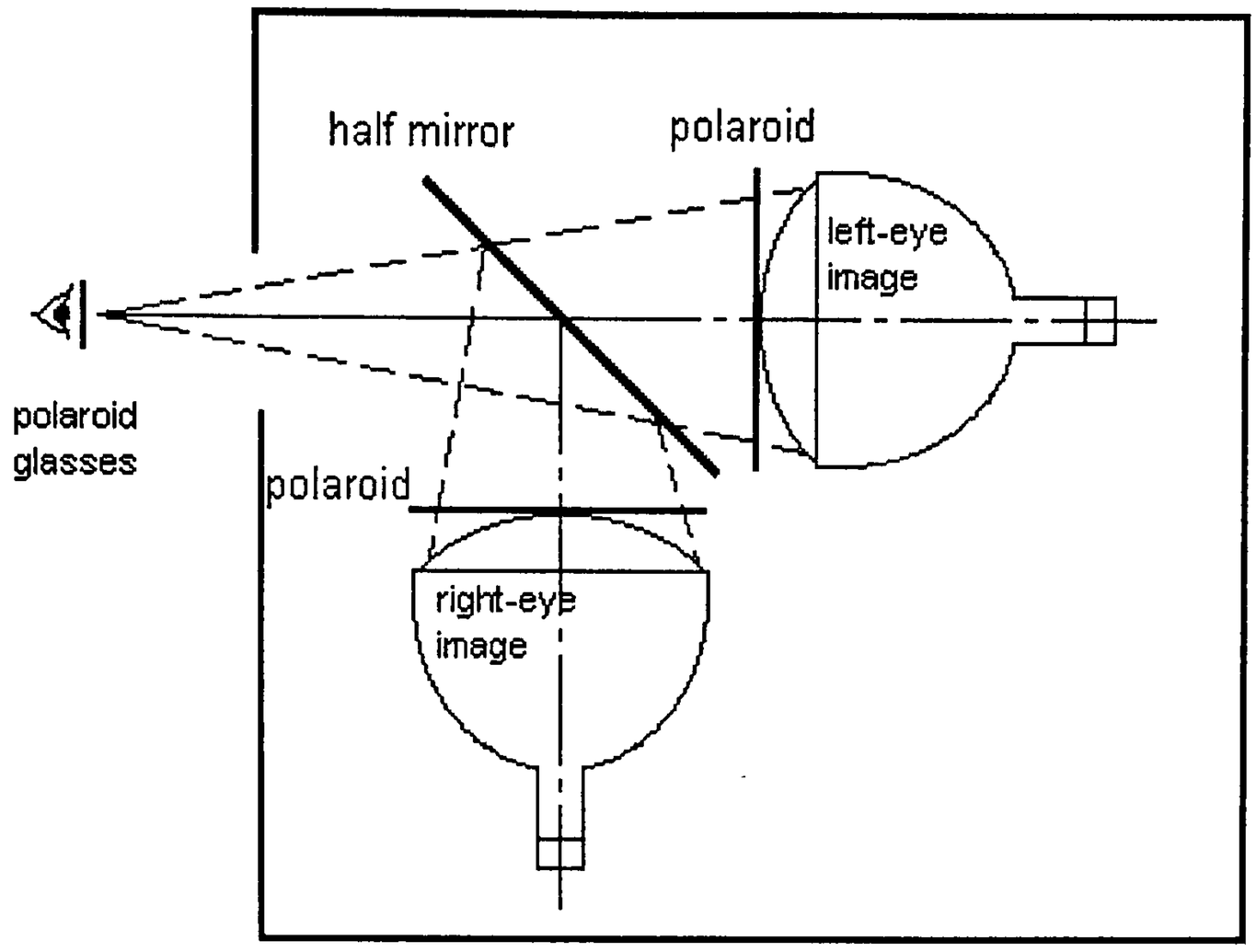


Figure 2.2: A binocular television system using two cathode-ray tubes ([74]).

In normal unaided sight, the lines of sight converge and this convergence gives the depth perception. At the same time, accommodation of the crystalline lens of the eye takes place such that accommodation and convergence agree with each other: when the eyes focus on an object (accommodation), the lines of sight intersect at the object (convergence). In the case of viewing a stereo pair, the lines of sight are essentially parallel or focused at infinity, whilst the eyes are focusing at close range ([84]). This contradicts the nature of the human eye physiology and may cause eye strain.

All stereo pair displays need an image-selective device between the viewer and the screen, bringing forth a series of drawbacks related not only to the comfort of the viewer, but also to the channelling methods themselves ([64]). The displays which present the left and right components of a stereo pair using colour coding produce eye strain due to the rivalry between the two retinas and suffer from colour rendering problems ([64]). Polarising or time division shuttering spectacles produce similar eye strains as light reaching the eyes is attenuated. In the case of electro-optic mechanical shuttering, the two images may not be completely separated due to the relatively long image persistence, known as the 'afterimage' effect of the cathode ray tube ([64], [74]). These problems are reduced to a certain extent if liquid crystal shutters are used ([64]).

It is generally accepted that the restrictions imposed by viewing devices limit the large scale adoption of stereoscopic systems in applications involving audiences, like TV and movies. However, there are many entertainment and commercial systems, which use such a technology ([64]). Moreover, binocular systems are of special value in the medical and educational fields ([74]).

2.2.2 Multiview based systems: virtual studios

The state-of-the-art advances in stereo systems are the virtual reality systems, created by binocular or multiview images and aided by high speed computing algorithms, which can present spatial (binocular parallax) views within a scene according to the viewer's head position. The spatial effect is emphasised by the temporal parallax, i.e. change of the viewing position, which is continuous and unrestricted ([64]).

Virtual studios compose, in real time, live video with synthetic or natural imagery. The hardware and software used to achieve this visual medium make a virtual studio system, whilst the imagery constitutes a virtual set ([34]). The viewer is immersed in a virtual reconstruction of real world events and is able to select a viewing angle at view time and freely move throughout the virtual (or virtualized) event ([44]). A stereoviewing system allows the viewer to be immersed in a stereoscopic reconstruction of the virtual event. The virtual reality (VR) 3D-models are created either using CAD tools, or by synthesis from real-world images ([44], [66]).

A typical virtual(-ized) reality process involves three stages ([44]):

1. *Data (image) acquisition about the visual event.* Generally, data acquisition involves recording a scene from multiple positions using strategically located cameras. Moezzi et al. ([66]) used 17 cameras surrounding a stage, whilst Kanade et al. ([44]) used up to 51 cameras placed in the nodes of a hemispherical dome, and Tseng and Anastasiou captured multiple simultaneous views using a set of equidistantly aligned cameras with parallel axes ([91]).

2. *Recovery of the 3D-structure of the scene.* For recovering the 3D-scene, Kanade et al. and Tseng et al. used stereo methods to compute time-varying

depth maps which encode the scene depth of each image point from the corresponding initial views. A scene description is a combination of an image and an aligned depth map. Scene descriptions translate easily into computer graphics models, whose geometry is taken from the depth map and whose texture is provided by the image ([44]). Moezzi et al. concentrate on the accurate recovery of the 3D-shapes of dynamic or foreground objects with a volume occupancy method ([66]). The volume enclosing the scene is divided into voxels and an iterative volume intersection method is applied to decide if the voxel is occupied by the object or not.

3. *Generation of synthetic viewpoints.* Visual reconstruction from arbitrary viewpoints, even views for which no physical camera exists, is a very important phase of the virtual reality process. Kanade et al. used the depth maps from the closest camera to the viewer to render the scene and additional selected depth maps to handle the object occlusions and improve the quality of the synthesised image ([44]). Tseng and Anastasiou ([91]) generated virtual images by interpolating real views scanline by scanline using disparity information. Moezzi et al. ([66]) created 3D-models with fine polygons, each separately coloured, and used standard graphics libraries to render them.

In order to work properly and achieve accurate visual reproduction of real scenes, virtualized reality needs a large number of cameras to acquire the video sequences from various angles and a good synchronisation between them. Systems able to capture synchronous frames from multiple cameras have been reported ([68]).

These systems promise to have multiple applications in realistic training in a virtualized work space and true telepresence ([44]), but they are still too expensive and pose health risks (headaches, eye strains) to be used for three dimensional television ([64]).

2.2.3 Depth extraction in stereoscopy

The final aim of the present work is to derive a depth extraction method for integral imaging. It is therefore useful to dwell upon the way depth estimation has been approached in stereoscopy.

The simplest stereoscopic setup is presented in Figure 2.3. It comprises two cameras of focal length f with parallel optical axes, separated by a distance B. The line connecting the camera lens centres, called *baseline*, is orthogonal to the optical axes of the system. A global Cartesian system having the origin O in the middle of the baseline and the x-axis coinciding with the baseline. An object point (x, y, z) has image coordinates (x'_l, y'_l) in the left and (x'_r, y'_r) in the right image plane. By similar triangles, it is obtained that the depth of the object point is ([62]):

$$z = \frac{B\,f}{x'_l - x'_r} = \frac{B\,f}{d}\,, \tag{2.1}$$

where $d = x'_l - x'_r$ is called the *disparity* between the corresponding left and right images of the physical point. Equation (2.1) shows that theoretically the depth of a point is determined when two corresponding images are known. However, there are a few difficulties that accompany depth extraction from a stereo pair([62]). Near objects can be accurately recovered, but extracting the depth of far away objects proves difficult: disparity (d) can only be measured in pixels and this approximation causes big distortions when z has a large value. Moreover, because the disparity is proportional to the baseline B, any error in the evaluation of the disparity increases when the baseline increases. It results in a subsequent depth error. However, if the baseline becomes large, it is also difficult to correlate the two images and match the corresponding points. This is known as the *stereo correspondence problem* ([62]).

Ohta and Kanade proposed a stereo matching scheme using the dynamic programming technique ([72]) applicable to stereo images rectified such that the epipolar lines are horizontal scanlines. The algorithm uses edge-delimited intervals as elements to be matched and employs two types of searches: pairs of corresponding points are searched within the same scanline (intra-scanline search); another search is performed inter-scanline for possible correspondences of connected edges in the right and left images, which insures the consistency of the results across the images.

Another successful approach of the problem uses redundant images, which can come from multiple camera systems or fine motion image sequences, for increasing precision at depth estimation ([63], [76]). Matthies and Okutomi ([63]) modelled the depth map as a random field and employed area based image similarity measure-

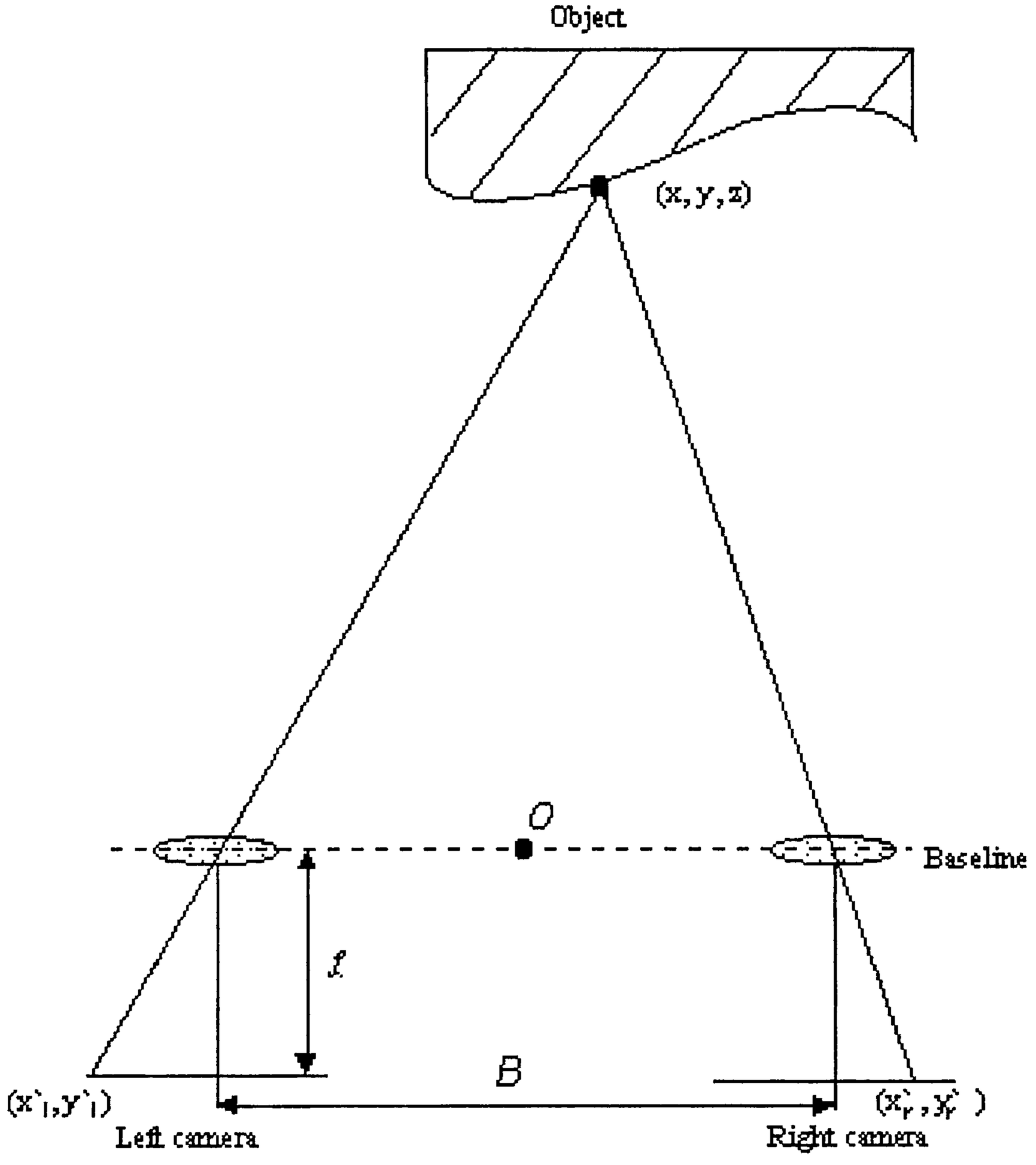


Figure 2.3: A stereo imaging system. ([62]).

ments, formulating the matching problem in statistical terms. The sensing strategy employs fine camera motion parallel to the baseline to acquire a narrow-baseline image pair. This initial image pair provides depth estimates which are used to constrain matching of a wide-baseline image pair ([63]). Okutomi and Kanade ([76]) derived a stereo matching method that uses multiple stereo pairs with various baselines generated by lateral displacement of a camera to obtain precise distance estimates and remove ambiguities. For each stereo pair, the sum of squared-differences (SSD) of the pixel intensity values is computed. The sums obtained for the individual stereo pairs are then added up to obtain the sum of SSDs (SSSD). The authors proved that SSSD exhibits a unique and clear minimum at the correct matching position, and that the method is able to eliminate false matches ([76]).

The final outcome of these methods is the depth of each point represented in the 2D-images. This information is used to produce the *depth map* of the space. The depth map is a two-dimensional array whose x and y coordinates correspond to the rows and columns in the 2D-picture, and the corresponding depth readings (z values) are stored in the array elements (pixels). It is like a grey scale image in which the intensity information is replaced with the depth information ([62]).

2.3 Holography

The principles of holography were discovered by Dennis Gabor in 1948 ([32], [33]), when he was studying the possibility of increasing the resolution of electron microscopes. To visualise the effectiveness of the method, he used a light beam instead of an electron beam, and created the first hologram ever made ([78]).

In his experiments, he considered a coherent light bundle collimated at a pinpoint and a tiny, transparent object placed in the diverging light bundle ([32], [74]) as shown in Figure 2.4 ([74]). Then the undisturbed light and the light scattered by the object, both emerging from a coherent light source, interfere with each other. The resultant interference fringes are recorded upon a photographic plate ([74]).

The initial system proposed by Gabor ([32], [33]) had the drawback of being able to reproduce only transparent objects and of obtaining the conjugate image in the same direction as the desired image. To remedy these difficulties, Leith and Upat-

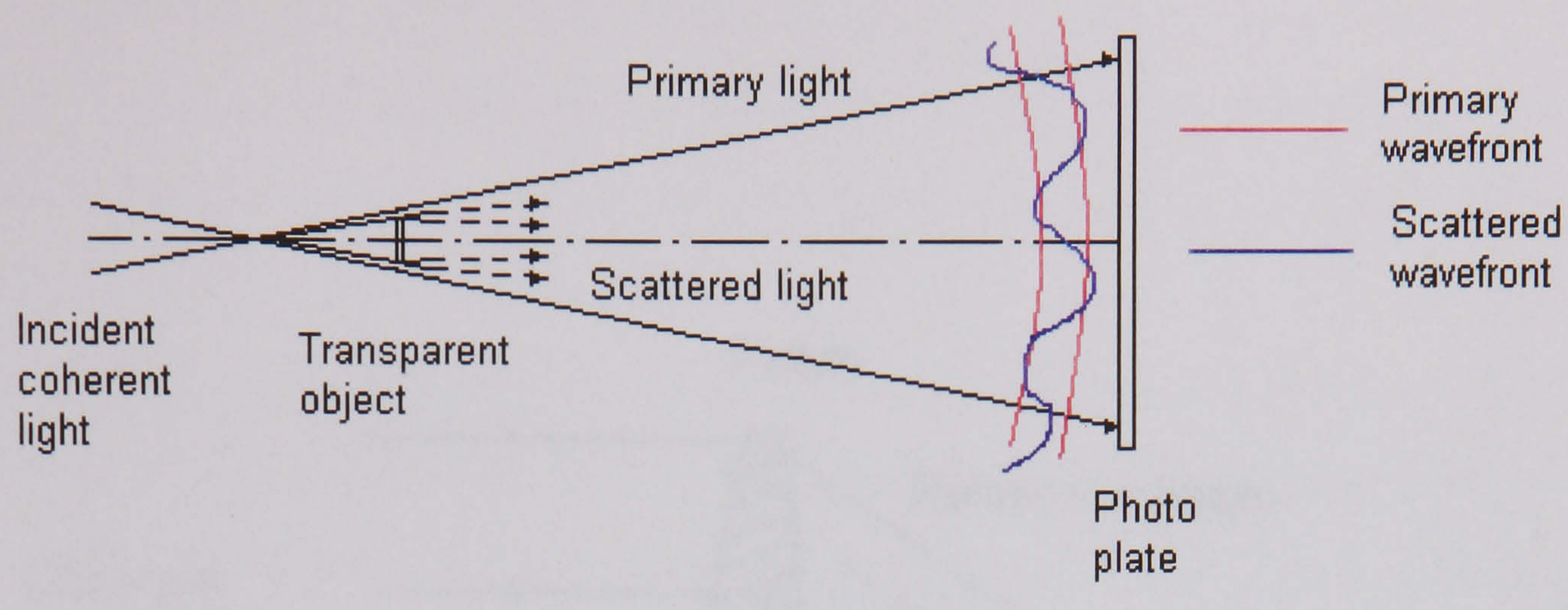


Figure 2.4: Foundation of holography: optical representation of Gabor's wavefront reconstruction method ([74]).

nieks ([49], [50], [51]) proposed a so-called two-beam method. Their arrangement, shown in Figure 2.5(a) ([74]) for the case of a transparent object, uses two light beams originating from a single coherent light source. The first beam is transmitted through the object and scattered before reaching the photographic plate. It is modulated according to the physical characteristics and dimensions of the object. The second beam, called 'reference beam', is refracted through a prism and illuminates the photographic plate as a plane wave. ([74], [78]). The resultant hologram contains the fringes produced by the interference of the two beams. The average spacing between two fringes is $d = \lambda / \sin\theta$, where λ is the wavelength and θ is the incidence angle of the reference beam (Figure 2.5(a)). The spacing and density of the fringes are modulated according to the scattering of the object beam ([74]). A system governed by the same principles, but additionally employing a set of mirrors, is used by Leith and Upatnieks to produce holograms of solid objects ([49], [50], [51]).

At reconstruction, a beam having usually the same wavefront used at recording illuminates the hologram ([74]). As a result of the diffraction upon the hologram, three light components emerge (Figure 2.5(b), [74]):

1. Bundle B_1 has the same wavefront as the recorded object wave and produces the direct image, which is virtual.

2. Bundle B_2 is the undisturbed light.

3. Bundle B_3 produces the conjugate image, which is real.

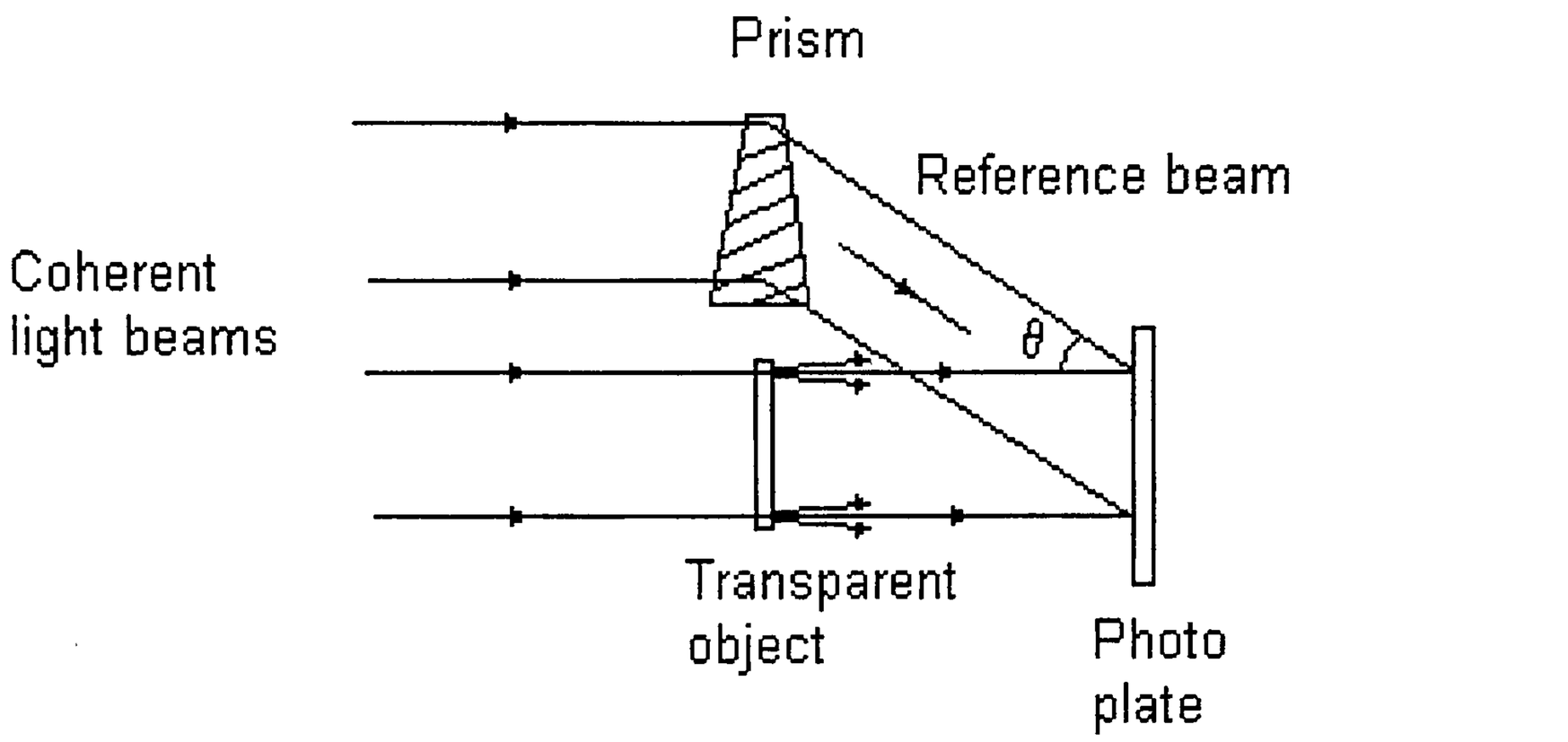


Figure 2.5: Leith-Upatnieks two-beam holography method: (a) hologram recording; (b) reconstruction from a hologram.

In most cases, the object will reconstruct to its original size, regardless of the size of the plate, and at the same distance it was when the hologram was made. The front of the object appears closer, the back farther away and all the points in between are filled in accordingly ([78]).

A hologram is very different from an ordinary photograph. A photograph records the intensity differences on the object which form a pattern that is ultimately recognised as the object photographed. A point to point correspondence exists between the object and the recording. In holography, interference fringes are recorded, and direction and phase of light are thus preserved and coded ([78]). In this way, each small portion of the hologram contains information about the whole object, and, if the hologram is broken or cut, each piece would be able to reconstruct the recorded image. This is because the light bouncing from each point on the object is not focused to a point on the film, but is allowed to spread out through the space between the object and the film, thus covering a large portion of the film as if each point were a spray of light each with a certain angle of divergence. Hence, every point is coded into a large area of the hologram. When a hologram is cut, each piece will still hold the whole image, unchanged in size or position, but with a changed viewing angle range ([78]).

The hologram, i.e. the medium that contains all the information, is a high contrast, very fine grain, black and white photographic film. The film designed for holography has to be capable of very high resolution, up to 3000 lines/mm. This high resolution is required by the fact that the grain has to be comparable as order of magnification with the wavelength of the light used ([78]).

The prevalent practical problem in producing a hologram is object movement. The slightest movement of the object completely obliterates the image. This is due to the fact that the interference fringes are separated only by a few wavelengths of light. When the object moves, the interference pattern changes at its turn, and a modification larger than a fraction of a wavelength has fatal results. This can be prevented either by completely isolating the object from all sources of movement, or by making the exposure very short. A quick exposure, around a billionth of a second, is possible only by using a very costly pulse ruby laser [78]. This laser system, developed in 1960 by T. Maiman from Hughes Aircraft Corporation ([38]),

emits a very powerful burst of light that lasts only a few nanoseconds and makes it possible to produce holograms of high-speed events and of living subjects ([38]).

From this discussion, several advantages and disadvantages of holography can be derived ([74], [78]):

- Features of holography ([74]):
  1. The 3D sensation given by holograms is satisfactory and based on binocular parallax, convergence and accommodation.
  2. It is a practically invulnerable recording scheme, in the sense that it can be scratched or cut without significant loss of reconstruction quality. This is due to the fact that any part of the hologram contains information about the whole object scene.
  3. It allows the storage of a huge amount of data on a limited physical support.
- Disadvantages of holography ([74]):
  1. Holograms are recorded in dark rooms.
  2. Stillness of the object is essential.
  3. Humans usually cannot be subjects of holograms. They cannot stay still for the exposure, and, moreover, the laser can damage the eyes.
  4. Ordinary holograms reproduce only monochromatic images.
  5. The efficiency of the reconstruction process is low.
  6. High resolution recording materials are very expensive.
  7. The amount of data recorded on a hologram is tremendous, much of it being redundant. It is not therefore an economical way to store and transmit 3D information.

Vigorous efforts have been made to overcome the drawbacks of holography, and significant progress has been reported in the direction of recording colour and animated holograms ([74]). However, due to the huge amount of data and health risks involved, holography is not considered a practical 3D-imaging system for movie and

television applications ([74]). However, the technique has numerous applications ranging from supermarket bar code scanners, telephone cards with embossed holograms, to archival recording of valuables or fragile museum artefacts ([38]).

2.4 Integral Imaging: a historical overview

Integral imaging is a technique able to create volumetric images through optical means and record them upon a single, 2D photographic plate. At replay, the optical models are reproduced three dimensionally and do not change as the viewer moves away from or towards them ([74], [12]). The principles of integral imaging were introduced in 1908 by G. Lippmann, who proposed the use of microlens arrays to record and replay 3D-images. The object is imaged through an array of microlenses and recorded on a photographic plate placed behind the microlens array. The resultant image contains therefore an encoding of the scene captured from numerous viewing points, which explains the multi-directional parallax observed at replay. After processing, the photographic transparency is overlayed with a microlens array similar to that used at recording and illuminated from the rear. In this way, a full parallax volumetric optical model of the object, which has the shape and colour of the original object, is reproduced by the intersecting ray bundles emerging from each of the lenslets. If the observer changes viewpoint, different directional beams enter the eyes presenting a continuously changing aspect of the scene ([64]).

The operation of Lippmann's system is depicted in Figure 2.6. It is noticed that the points of the objects situated at the front of the object are replayed further than those situated at the back of the object. An inversion in depth has therefore resulted and the image obtained is called *pseudoscopic* ([74]). H.E. Ives was the first to propose a method of curing this problem such that the replayed image is *orthoscopic*, i.e. in correct depth ([40]). His solution was to take a second integral photograph of the reconstructed image using a similar microlens array, as shown in Figure 2.7. This second photograph is the pseudoscopic inversion of the pseudoscopic image of the object, therefore it replays orthoscopically, as desired ([40], [74]). Ives' solution is not used in the present integral imaging systems because taking a second photograph of an integral image produces degradation in the resulting orthoscopic

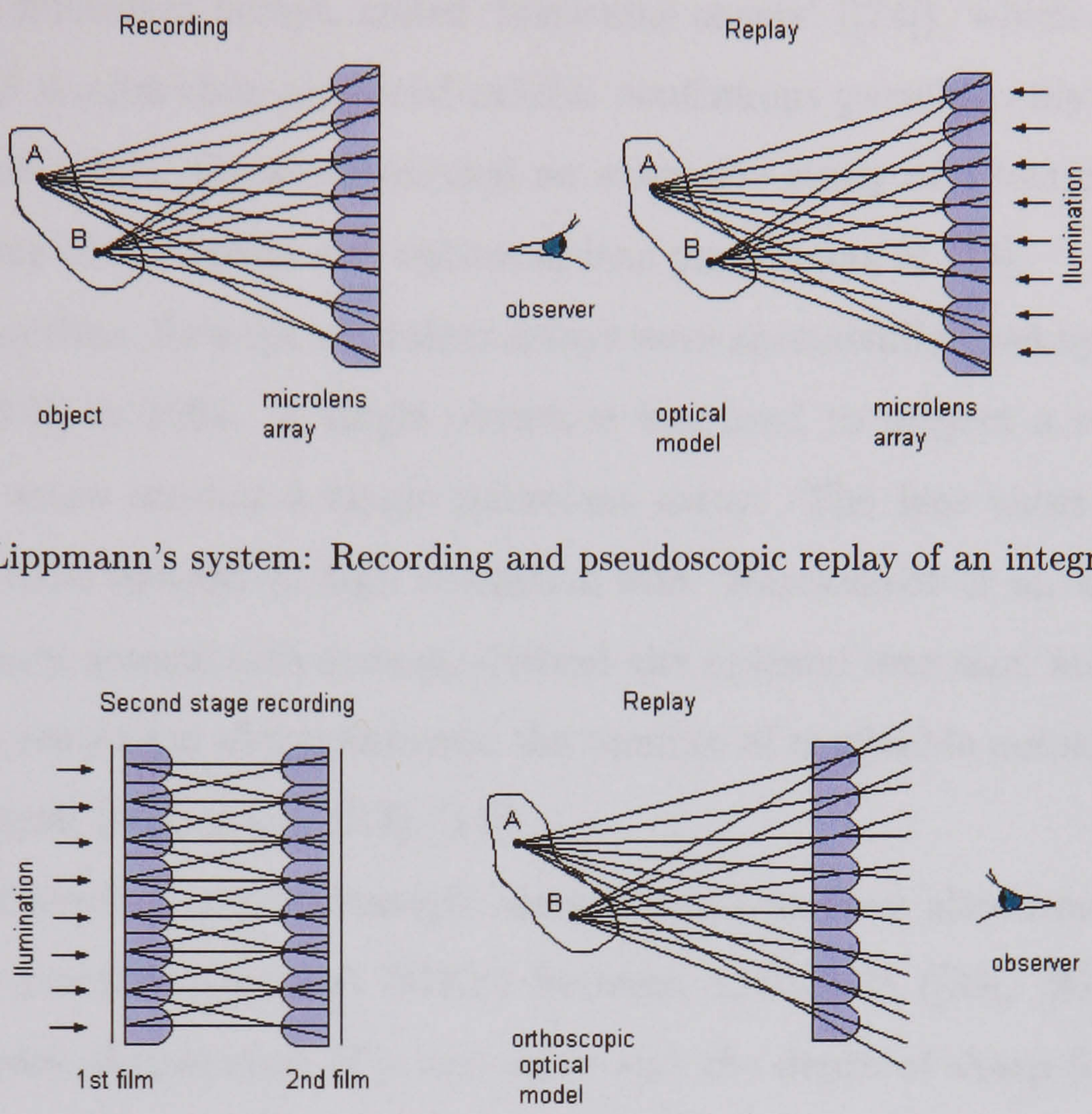


Figure 2.6: Lippmann's system: Recording and pseudoscopic replay of an integral photography ([12]).

Figure 2.7: Ives' two phase recording of an integral photography and orthoscopic replay ([12]).

integral image, but it has a tremendous theoretical importance and is the foundation of all further inversion solutions ([19], [20]).

Lippmann's idea of 3D-integral imaging was very much ahead the technological level of its time such that the development of integral photography was for a long while hampered by the lack of quality microlens arrays. The first experiment with integral photography is said to have been performed in Russia in 1911 by A.P. Sokolov ([93], [74]), who used a 'pinhole sheet' instead of a microlens array to image the filament of an incandescent lamp and obtain definite depth sensation. According to Valyus ([93]), the first integral photography obtained with a lens array was produced in 1948 by S.P. Ivanov and L.V. Akimakina. The lens array contained two million lenslets, each having a diameter of 0.3 mm and a focal length of 0.5 mm.

Post-war technological progress in precision machining and plastic materials allowed the manufacturing of better optical quality microlens arrays and a wider freedom of material choice. Together with the progress in photographic and printing technologies, it facilitated a renewed interest of the scientists in integral imaging ([74]). A simplified version of integral imaging has been considered: it uses cylin-

drical surface microlens arrays, called 'lenticular sheets' ([74]), which are easier to fabricate. The images thus produced exhibit continuous parallax only on the horizontal direction ([74]). Okoshi presented an extensive study of a lenticular integral system in terms of resolution and optimum lens parameters in [74].

At the same time, fly's-eye microlens arrays were successfully used by Burckhardt et al. ([13], [14]) in 1962. A single objective was used to project a reduced scale image of the scene around a single microlens array. The lens sheet used was a layer of glass balls bonded to high resolution film. Burckhardt et al. carried out a parametric study around this system, derived the optimal lens size, and calculated the maximum resolution obtainable and the number of resolvable points required to record an integral photograph ([13], [14]).

Parametric studies on pseudoscopic integral imaging have also been carried out by a Russian research group at NIKFI between 1970-1986 ([39], [23], [24], [25], [81]). The modes of operation of a lens array and the depth of sharp focusing have been analysed with respect to the camera parameters ([25]). A so called 'real index method' was proposed to make measurements of objects recorded in integral images. The method consisted of superimposing a physical index mark, in the form of the luminous tip of an optical fiber, on points of an integral (spatial) image of the object and deriving the size of the object from the coordinates evaluated for the marked points ([23]).

The latter decade brought forth a novel integral imaging arrangement, termed a 'two tier integral system' proposed by Davis and McCormick ([19], [20], [21]) and developed within the 3D-Imaging Group at De Montfort University. The most innovative feature of this system is its capability of single-step orthoscopic recording, via a double depth reversal process ([19], [12]). The arrangement employs a transmission element made up by a pair of back to back microlens arrays which produces spatial inversion. The result of the imaging process is a full colour, 1:1 orthoscopic volumetric optical model of the object, which is then recorded on photographic film placed behind a microlens array. The system will be extensively discussed in Chapters 3-4, where a mathematical model will be presented.

Integral imaging resembles holography in the sense that it is recorded on a two dimensional physical medium and that it exhibits continuous parallax at replay.

However, the advantages of integral imaging over ordinary holography rely in the use of incoherent light at both recording and replay, in reproducing full natural colour ([64]) and the ability of recording moving as well as still pictures. In spite of not achieving the high resolution of holography, integral imaging is recommended as a more viable solution for domestic 3D television ([75], [64], [12]). When comparing with existing multiview stereo systems, the outstanding feature of integral imaging is that it produces and *displays optical* 3D-images, creating a natural depth perception ([74]). Therefore, it alleviates the deficiencies associated with the health risks posed by stereoscopic systems. Virtual reality produced by stereoscopic systems is currently no substitute for the naturally perceived 3D world, therefore integral imaging is a good alternative method for providing true spatial imaging for entertainment, scientific, and technological purposes in the future ([64]).

2.5 Conclusions

An overview of the main directions in the development of 3D imaging has been presented. Integral imaging is a promising technique whose application in 3D television has been envisaged for the coming decade ([74], [75], [64], [73]).

Depth extraction methods used in stereoscopy have also been discussed. They are based on stereo matching and disparity estimation. Their relevance to the corresponding problem in integral imaging is however limited due to the particularities that separate the two systems.

The present work studies integral imaging from the perspective of decoding the depth information stored in a planar recording of a 3D-integral image. A mathematical model of the operation of the two tier integral system is presented in the following two Chapters, which leads to an understanding of the image formation and recording process. This model reveals the mechanism of depth encoding in a recorded integral imaging.

Chapter 3

Mathematical Model of the Unidirectional Integral Imaging System

3.1 Introduction

The present chapter tackles the problem of image formation and recording in the case of the 3D-unidirectional (lenticular) integral camera system. The problem is of tremendous importance as it brings an insight into the manner spatial, directional, and, subsequently, depth information is spread across the planar recording of an integral image.

After a detailed description of the integral two tier camera system and its operation, the analysis is focused on two main aspects of the involved optical process:

1. The geometry of the image formation: image coordinate computation, disparity, imaging fields, number of recorded intensity distributions for a single point.

2. The physical optics aspects of the process: point spread function and variation with respect to depth.

Depth influence onto image formation and recording is extensively analysed in section 3.5.

3.2 Integral Imaging Camera System

The development of 3D-imaging systems has a long history and involves a wide range of research fields. Integral photography was pioneered by Lippmann [53] (1908), who used microlens arrays to record on photographic film, and, subsequently, to replay integral images in three dimensions. Since 1908, when it was first reported, the integral photography technique has been improved as a result of theoretical studies, technical innovations of the optical systems, and progress in microlens manufacturing to such a level that it is now a viable alternative to other autostereoscopic systems.

The optics of an advanced form of integral imaging system – the two tier optical network – in which a true 3D optical reconstruction of a scene is transferred to the capture device, has been developed at De Montfort University and described in detail by Davies and McCormick ([19]-[21]). The optical arrangement relies on the use of two macrolens arrays (MA1 and MA2 in Figure 3.1) placed equidistantly behind and in front of an autocollimating transmission screen (ATS in Figure 3.1). The recording plane - a photographic plate whose position coincides with the focal plane of a microlens array (RA) is placed in the image space, within the volumetric optical model of the scene.

To enable accurate descriptions of the integral system, a number of terms will be used in order to express the type of information distribution associated with an imaged scene: an *image* is defined to be a point by point reconstruction of the object space; an *optical model* is considered to be formed by intersecting modulated ray bundles; a *voxel* represents a resolvable point in space; the final recorded data is a *sampled set of disparate micro-intensity distributions* ([56]).

The autocollimating transmission screen ATS is made up of two microlens arrays placed back to back and separated by their joint focal distance. It operates by refracting the transmitted ray at an equal angle to the angle of incidence of the impinging ray. In this way, every object point in front of the screen has an equidistant optical model behind the screen. Consequently, the ATS reverses the spatial sense of the imaged scene and enables the output macroarray to present a pseudoscopic image for recording such that the result on replay is a true 3D orthoscopic reproduction of the scene. The autocollimating transmission screen is able to

transfer a true optical reconstruction of an object or scene to a different location. The reconstruction contains all the spatial information about the object or scene with unit magnification. In order to achieve good lateral resolution and smooth continuous parallax in holoscopic imaging, it is necessary to use small (90-125 μm) micro-optical elements (square, circular or hexagonal based) in the screen ([19]).

The drawback of an optical arrangement which uses only the ATS for generating integral images comes from commercially available fine pitch microlens arrays having a very short range (approximately 10 mm) over which they reconstruct a sharp image. To overcome this limitation, the object space is imaged close to the fine pitch microlens surface using an array of relatively large relay lenses (MA1). In this way, an extended scene is compressed into the focal range of the relay lens image and this image can be arranged to occur at or near the screen plane and be within its imaging range. As each microlens receives and returns images of the relay lens from different directions, this combination retains a high sampling point number while greatly extending the range over which sharp definition can be maintained. The second macrolens array (MA2) has the role of integrating the intermediary optical models of the imaged scene. In this way, continuity of parallax throughout the viewing angle of the device is achieved. This arrangement has been named *"the two tier optical network"* ([21]).

The optical process can be summarised as it follows ([21], [56]):

- The object is imaged by the *input macrolenses* (MA1) which transmit compressed transposed images that occur at or near the central double microlens screen (ATS).

- The *autocollimating transmission screen* (ATS) inverts the optical sense of each intermediary image (Figure 3.2), and, simultaneously, presents these spatially reversed 3D-optical models to the corresponding output macrolenses.

- The *output macrolenses* (MA2) transfer and retranspose these optical models to their correct spatial location. The final integrated image, formed by superposed optical models projected by the second macrolens array, is a true full colour 3D optical 1:1 reconstruction of the object.

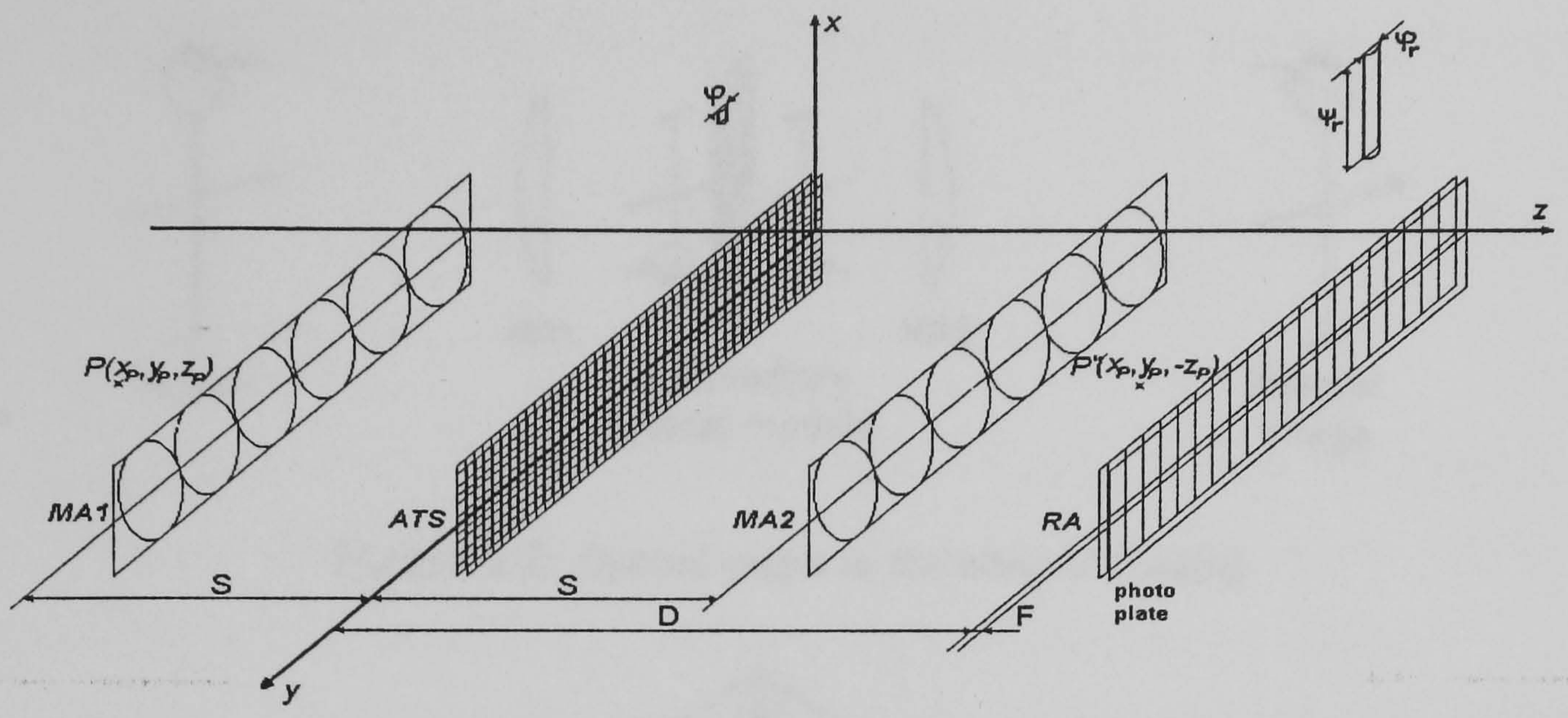


Figure 3.1: 3D-lenticular integral imaging camera system – two tier optical network.

- This optical reconstruction can be recorded as a *sampled data set* possessing either:

  - horizontal parallax – if semicylindrical microlenses have been used in the recording array RA,

  or

  -full parallax in all directions – in the case of an arrangement using circular, hexagonal or square based microlenses in the recording array.

As mentioned, the integral image is recorded in the focal plane of a recording microlens array as a planar sampled data set which contains multi-directional 3D information related to the object space. Each microlens of the recording array samples a fractional part of the scene, and, at the same time, many microlenses record directional information of the scene from different viewing angles. Therefore, parallax information about any particular point is spread over the image plane. Redisplay of the full spatial model as a real 3D image can be effected by overlaying the sampled data set by an integral decoding element.

The two tier optical network generates an optical model with unit magnification. For practical movie, TV, and compact camera systems, this image must be reduced in scale such that an extended scene can be compressed into the required frame size for electronic or photographic recording. This can be achieved by replacing the output macrolens array with a "segmented lens" (the lenses employed in the output array have a different focal distance and a different axis from those of the input

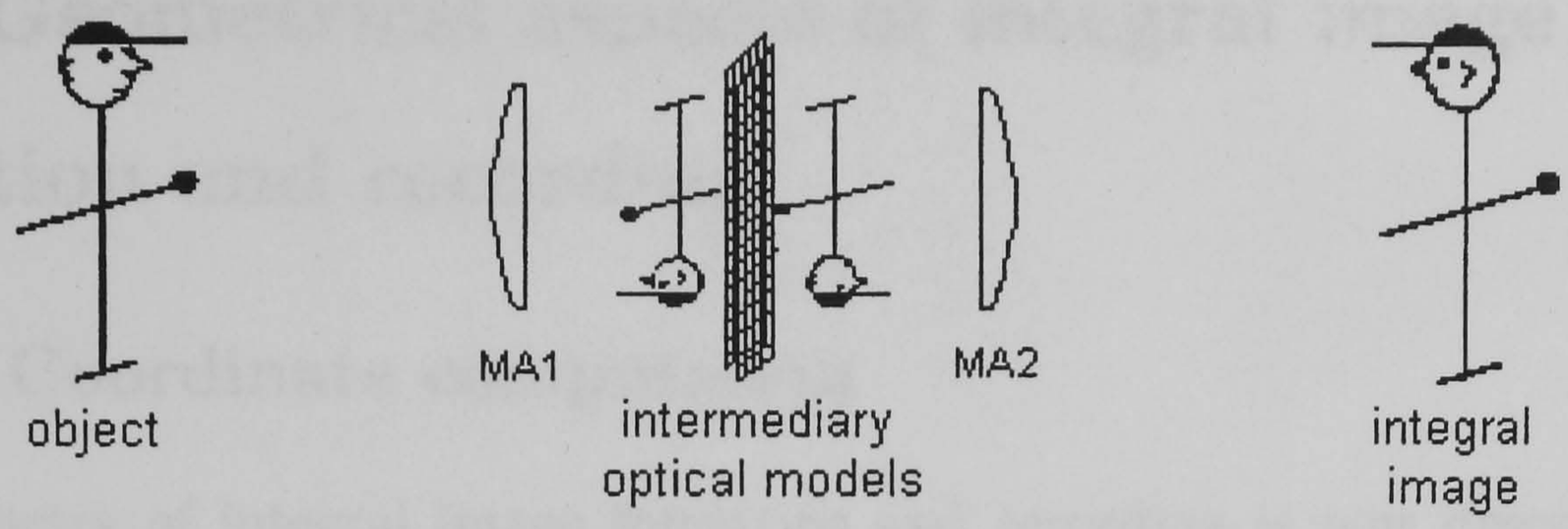


Figure 3.2: Optical stages in the integral imaging

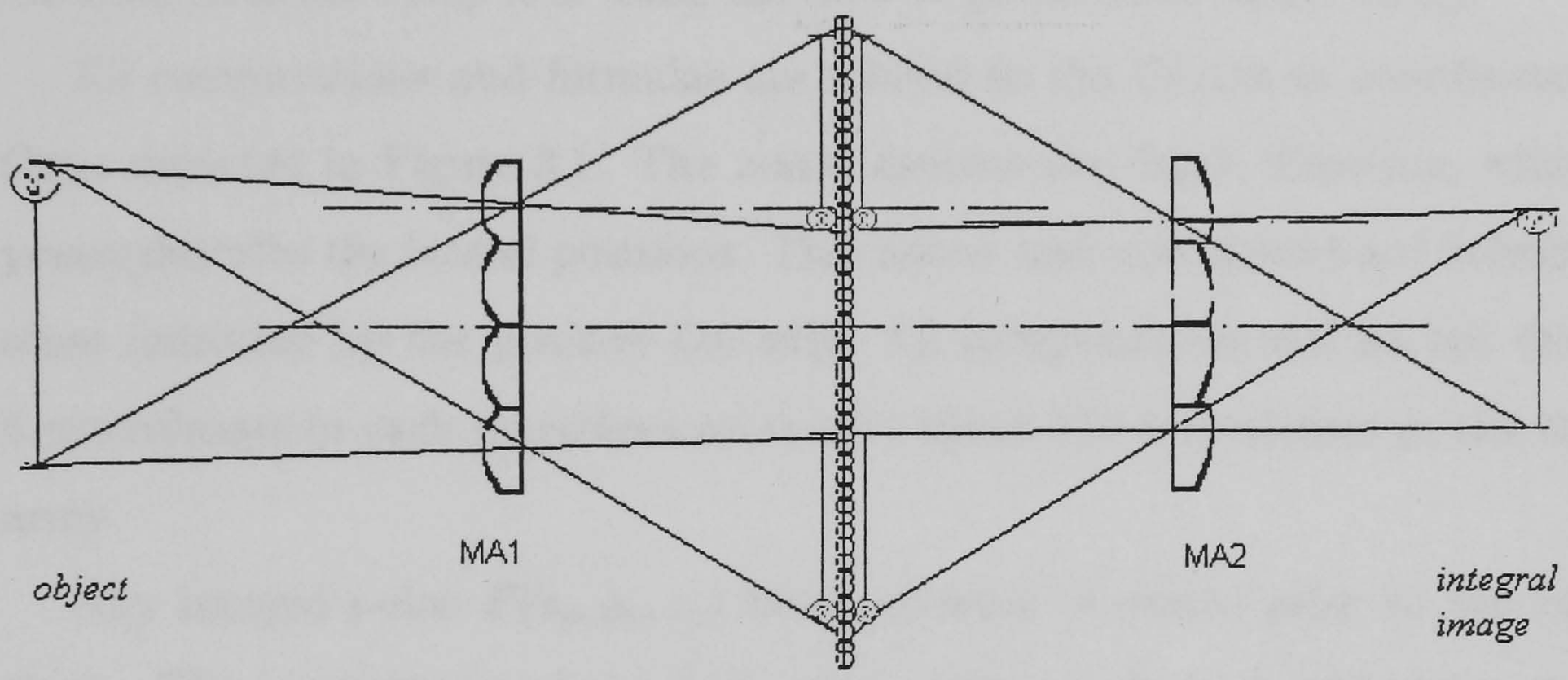


Figure 3.3: Segmented lens two tier camera system.

lenses, therefore the scene can be optically reconstructed at a smaller scale) - Figure 3.3 ([21]).

The present chapter concentrates on the optical process involved using the two tier optical network camera system of Figure 3.1 for the case where the ATS comprises square based spherical surface lenslets and the recording array is made up of identical cylindrical surface microlenses. This system is known as the 3D-unidirectional (or lenticular) integral imaging system ([74], [19]) and produces 3D images containing continuous horizontal parallax.

A mathematical model of the 3D-lenticular integral camera system based on the two tier optical network has been developed for the first time in order to understand how the spatial information (especially the point depth) is encoded in the final recorded image ([56]). Both geometrical and physical aspects of the optical process are analysed in detail.

3.3 Geometrical aspects of integral image formation and recording

3.3.1 Coordinate computation

The geometry of integral image formation and recording is now described in full detail. The coordinates of the intermediary and final optical model, and of the recorded data are computed using the laws of geometrical optics ([56]).

All computations and formulae are related to the Cartesian coordinate system $Oxyz$ depicted in Figure 3.1. The z-axis denotes the depth direction, while the x, y-axes describe the lateral positions. The macro- and microlenses are indexed in the sense indicated by the positive Oy axis. All computations carried out considered 6 macrolenses in each macrolens array and about 250 microlenses in the recording array.

Any imaged point $P(x_p, y_p, z_p)$ has a number of states prior to the recording stage. The macrolenses of the first array present a first set of images of P near the autocollimating transmission screen. These images are located at the points P_k^1 having coordinates ([56]):

$$\begin{aligned} x_k^1 &= \frac{x_P f_1}{f_1 + S + z_P} \\ y_k^1 &= \frac{b_k z_P + b_k S + y_P f_1}{f_1 + S + z_P} \\ z_k^1 &= \frac{z_P f_1 - z_P S - S^2}{f_1 + S + z_P}, \end{aligned} \tag{3.1}$$

for any macrolens k in the first array, where: f_1 is the focal length of the macrolenses, b_k denotes the y-coordinate of macrolens k 's centre, and S is the distance between a macrolens array and the central plane Oxy (see Figure 3.1 and Figure 3.4(a)).

These intermediary images become object points for the ATS. The light emerging from them, refracted and diffracted through the first microlens array, produces an intensity distribution in the common focal plane of the two microlens arrays; this intensity distribution is transmitted by the second microlens array through a symmetric optical process. As a result, the autocollimating transmission screen operates by refracting the transmitted ray at an equal angle of opposite sign to the

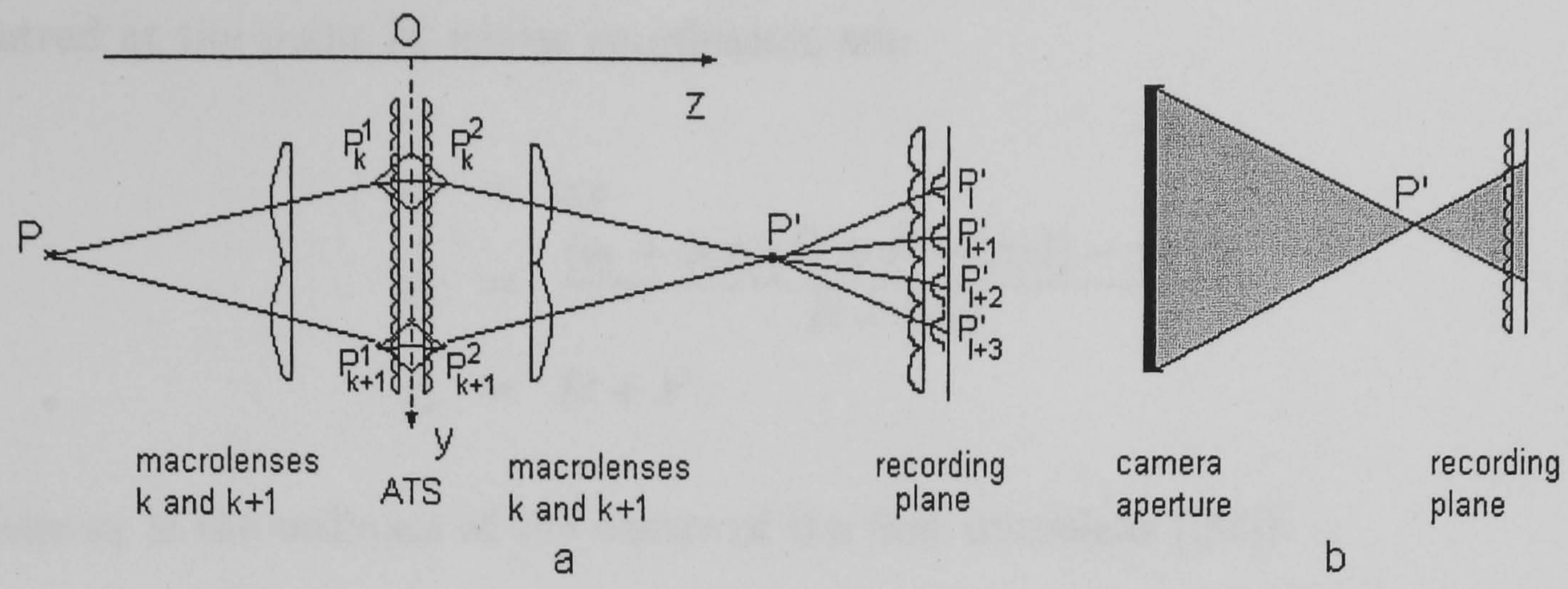


Figure 3.4: Transverse sections of the arrangement in figure 3.1 with a plane $x = x_P$: (a)Schematic representation of the intermediary, integral and recorded optical models of a physical point P; (b) The microlenses of the recording array that record a certain point P'

angle of incidence of the impinging ray. In this way, every "object" point P_k^1 in front of the screen has a quasi-equidistant voxel representation centred at the point $P_k^2(x_k^2, y_k^2, z_k^2)$ behind the screen, namely ([56]):

$$\begin{aligned} x_k^2 &= \frac{x_P f_1}{f_1 + S + z_P} &&= x_k^1 \\ y_k^2 &= \frac{b_k z_P + b_k S + y_P f_1}{f_1 + S + z_P} &&= y_k^1 \\ z_k^2 &= \frac{-z_P f_1 + z_P S + S^2}{f_1 + S + z_P} \end{aligned} \tag{3.2}$$

Each intermediary voxel P_k^2 is collected by the corresponding output macrolens k and imaged at the location $P'(x_P, y_P, -z_P)$ in space, which is the equal conjugate image location of point P with respect to the Oxy plane. It does not depend on the macrolens k, and therefore the result is an integration of directional information about P from all the macro- and microlenses that have imaged it ([56]).

The intensity distributions related to the integral image P' of P are recorded on a photographic plate that lies behind the recording microlens array containing microlenses of pitch ϕ_r, length ψ_r, and focal length F, situated at a distance D from the central Oxy plane - Figure 3.1. P' has recorded intensity distributions in a number of microlenses, the number being determined by the axial position of P'. If microlens k "sees" the point P', then the recorded data behind microlens k is

centred at the point P'_k whose coordinates are:

$$\begin{aligned} x'_k &= x_P \\ y'_k &= \frac{(c_1 + \phi_r k)(D + F - |z_P|) - y_P F}{D - |z_P|} \\ z'_k &= D + F\,, \end{aligned} \tag{3.3}$$

where c_1 is the ordinate of the centre of the first microlens ([56]).

3.3.2 Disparity and point position retrieval

The disparity between two recorded intensity distributions of the point P corresponding to microlenses k and j is ([56]):

$$d_{jk} = \frac{(D + F - |z_P|)|j - k|\phi_r}{D - |z_P|}\,. \tag{3.4}$$

This expression allows the recovery of the physical position of the point when two of its intensity distributions, centred at $P'_j(x'_j, y'_j, D + F)$ and $P'_k(x'_k, y'_k, D + F)$, are known (and, therefore, the disparity d_{jk} between them is known). Namely:

$$\begin{aligned} x_P &= x'_j = x'_k \\ y_P &= \frac{(c_1 + \phi_r j)d_{jk} - y'_j|j - k|\phi_r}{D - |z_P|} \\ z_P &= \frac{(D + F)|j - k|\phi_r - d_{jk}D}{d_{jk} - |j - k|\phi_r}\,. \end{aligned} \tag{3.5}$$

3.3.3 The imaging field of a microlens

Each microlens of the recording array samples a fractional part of the optical reconstruction of the scene and produces a recorded data set. The location of this data set on the photographic plate represents the imaging field of the microlens.

Considering a *pinhole approximation* model for a microlens k of the recording screen, the intersection between the imaging field and a certain plane $x = x_0$ is the segment $[A_k B_k]$ (Figure 3.5). The two ends of this segment are:

$$A_k\left(c_k\frac{D - S + F}{D - S}, D + F\right)$$

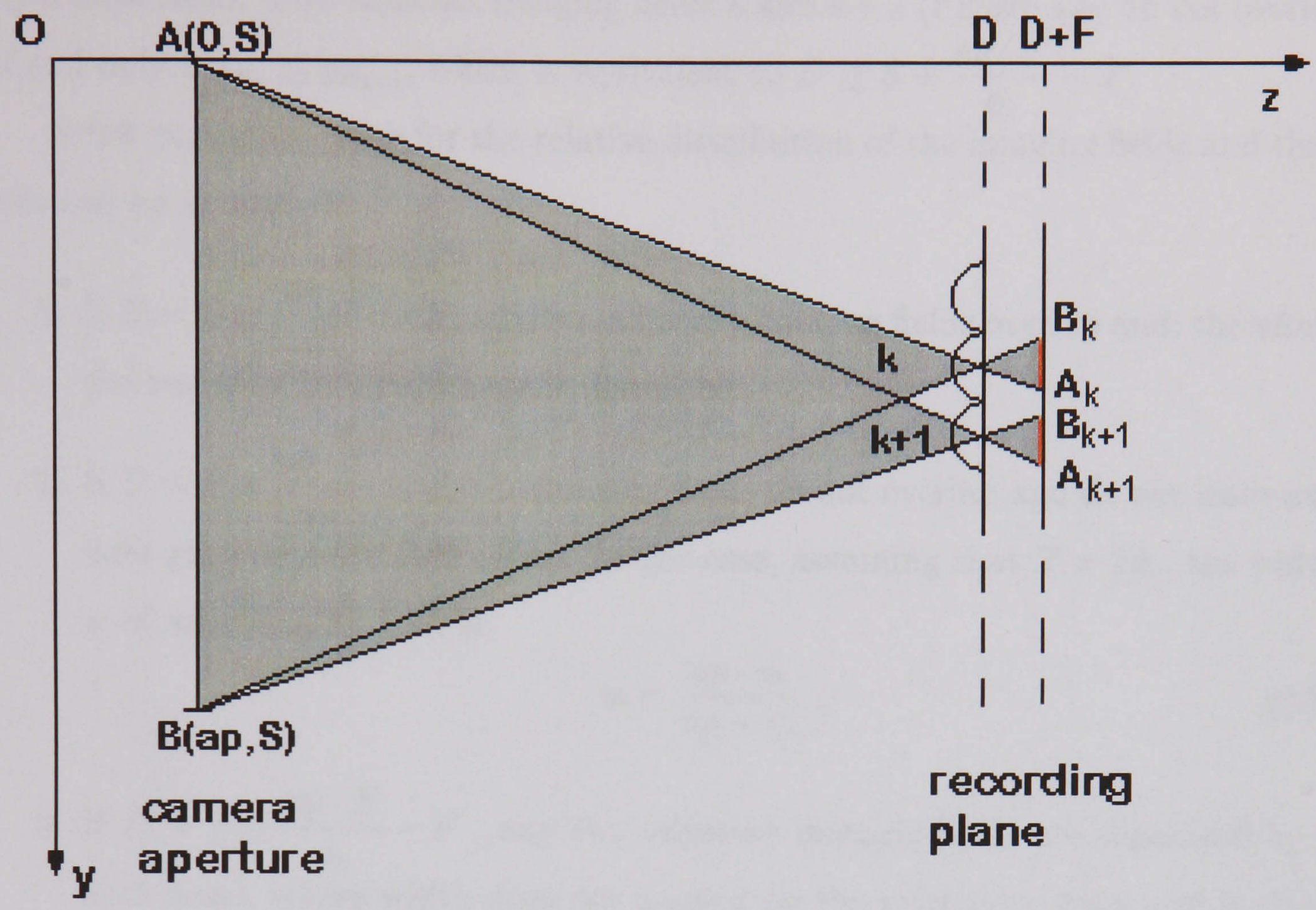


Figure 3.5: The imaging fields of the microlenses k and $k+1$ of the recording screen.

and

$$B_k\left(ap + (c_k - ap)\frac{D - S + F}{D - S}, D + F\right),$$

where $c_k = \dfrac{2k-1}{2}\phi_r$ is the ordinate of the centre of the microlens k. It follows that the width of the imaging field of microlens k is:

$$w(k) = \|A_k B_k\| = \frac{ap \cdot F}{D - S}. \tag{3.6}$$

It can be noticed that this expression does not depend on k, thus the imaging field has the same width for all the microlenses of the recording screen.

The centre of the imaging field (which is the middle of the segment $[A_k B_k]$) is

$$m(k) = c_k\frac{D - S + F}{D - S} - ap\frac{D - S + 2F}{2(D - S)}, \tag{3.7}$$

and, in general, does not coincide with c_k. Therefore, the imaging field centre is displaced from the corresponding microlens axis.

The imaging fields of two adjacent microlenses can overlap or can be separated

by a dark band. Two adjacent imaging fields k and $k+1$ (Figure 3.5) do not overlap if and only if $y_{A_k} \leq y_{B_{k+1}}$, which is equivalent to $D \geq S + \frac{ap \cdot F}{\phi_r} - F$.

A few particular cases for the relative distribution of the imaging fields and their size can be identified:

1. If $D < S + \frac{ap \cdot F}{\phi_r} - F$, any two adjacent imaging fields overlap and, therefore, the resulting integral image is disturbed.

2. If $D = S + \frac{ap \cdot F}{\phi_r} - F$, the imaging fields do not overlap and do not leave any dark gaps between each other. In this case, assuming that $F = 2\phi_r$, the width w of any imaging field is:
$$w = \frac{ap \cdot \phi_r}{ap - \phi_r}. \tag{3.8}$$

3. If $D > S + \frac{ap \cdot F}{\phi_r} - F$, any two adjacent imaging fields are separated by a dark band, whose width does not depend on the microlens index and is given by:
$$w_{bb} = \frac{\phi_r(D - S + F) - ap \cdot F}{D - S}. \tag{3.9}$$

The camera configuration can be set such that the width of the imaging field be equal to the microlens pitch ([12]). Assuming $F = 2\phi_r$, the condition $w(k) = \phi_r$ becomes: $D - S = 2ap$. This is a configuration which fulfills the condition of case 3. The dark bands have in this case a width $w_{bb} = \frac{\phi_r^2}{ap}$.

The imaging field modelling can be refined by introducing the influence of 2D-refraction ([12]). The three limit situations that occur at the image formation and determine the size of the imaging field are:

(a) the rays impinge on the front vertex of the lenslet k and generate a field $\mathcal{F}_1(k)$ (Figure 3.6).

(b) the rays cross in front of the lens k and impinge at the point of maximum sag and generate a field $\mathcal{F}_2(k)$ (Figure 3.6).

(c) the rays impinge on the point of maximum sag of the lens k before crossing and generate a field $\mathcal{F}_3(k)$ (Figure 3.6).

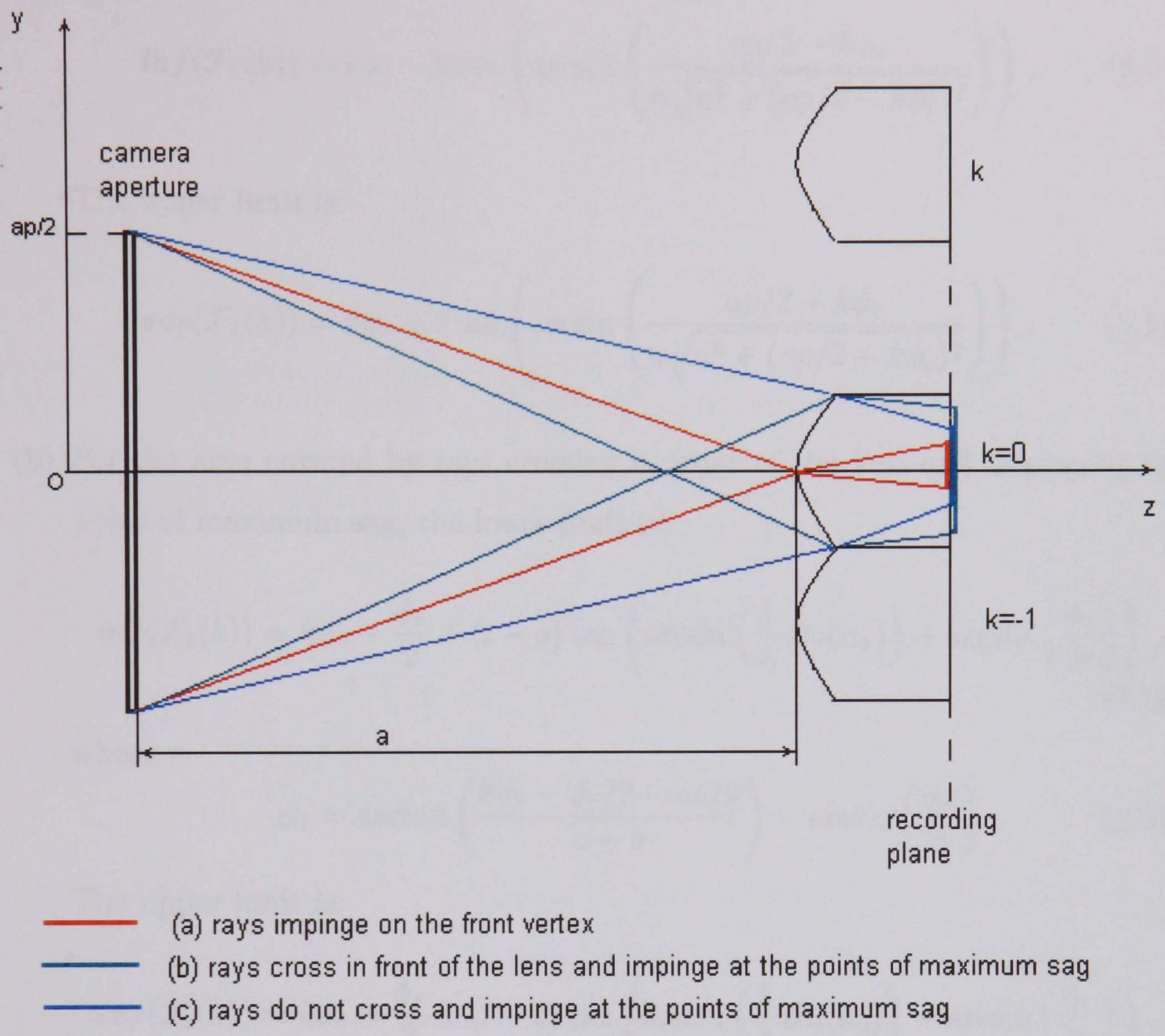


Figure 3.6: Schematic representation of the imaging fields distinguished when taking into account 2D cylindrical refraction and the contribution of the marginal rays.

The refraction imaging field has therefore three components whose upper and lower limits can be computed using the refraction law. A Cartesian Oyz frame with the centre situated in the centre of the camera aperture and the z-axis along the depth direction of the imaging system (Figure 3.6) has been considered. The microlenses are indexed symmetrically with respect to the central microlens, whose index is $k = 0$, as shown in Figure 3.6. The central microlens has the property that its centre and the centre of its imaging field are coaxial along Oz. All further calculations in this section are performed with reference to the frame in Figure 3.6.

The lower (inf) and upper (sup) limits for each of the three fields behind a certain microlens k are:

(a) The lower limit in the case when rays intersect at lens vertex is given by:

$$inf(\mathcal{F}_1(k)) = k\phi_r - t\tan\left(\arcsin\left(\frac{ap/2 - k\phi_r}{n\sqrt{a^2 + (ap/2 - k\phi_r)^2}}\right)\right). \tag{3.10}$$

The upper limit is:

$$sup(\mathcal{F}_1(k)) = k\phi_r + t\tan\left(\arcsin\left(\frac{ap/2 + k\phi_r}{n\sqrt{a^2 + (ap/2 + k\phi_r)^2}}\right)\right). \tag{3.11}$$

(b) For the area covered by rays crossing in front of the lens and impinging the point of maximum sag, the lower limit is:

$$inf(\mathcal{F}_2(k)) = k\phi_r - \frac{\phi_r}{2} + (t-s)\tan\left(\arcsin\left(\frac{1}{n}\sin(\alpha_1)\right) + \arcsin\left(\frac{\phi_r}{2r}\right)\right), \tag{3.12}$$

where

$$\alpha_1 = \arctan\left(\frac{k\phi_r - \phi_r/2 - ap/2}{a+s}\right) - \arcsin\left(\frac{\phi_r}{2r}\right) \tag{3.13}$$

The upper limit is:

$$sup(\mathcal{F}_2(k)) = k\phi_r + \frac{\phi_r}{2} + (t-s)\tan\left(\arcsin\left(\frac{1}{n}\sin(\alpha_2)\right) - \arcsin\left(\frac{\phi_r}{2r}\right)\right), \tag{3.14}$$

with

$$\alpha_2 = \arctan\left(\frac{k\phi_r + \phi_r/2 + ap/2}{a+s}\right) + \arcsin\left(\frac{\phi_r}{2r}\right) \tag{3.15}$$

(c) When rays enter the point of maximum sag without crossing, the lower limit of the field is:

$$inf(\mathcal{F}_3(k)) = k\phi_r + \frac{\phi_r}{2} + (t-s)\tan\left(\arcsin\left(\frac{1}{n}\sin(\beta_1)\right) - \arcsin\left(\frac{\phi_r}{2r}\right)\right), \tag{3.16}$$

with

$$\beta_1 = \arctan\left(\frac{k\phi_r + \phi_r/2 - ap/2}{a+s}\right) + \arcsin\left(\frac{\phi_r}{2r}\right) \tag{3.17}$$

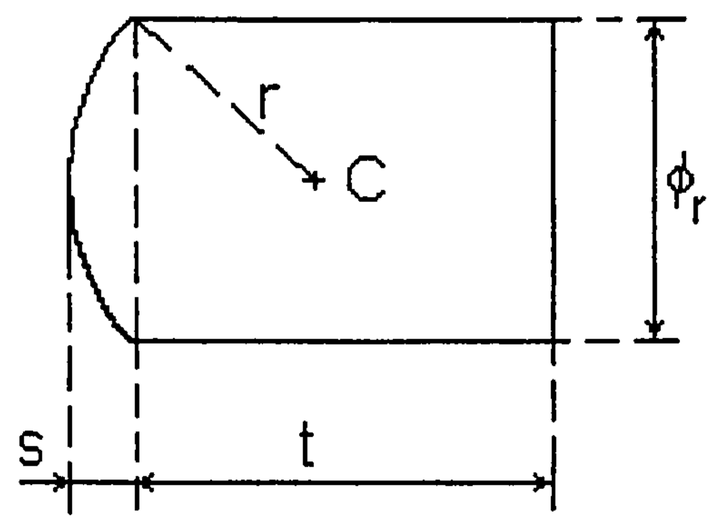


Figure 3.7: Lens parameters used in imaging field computations. C is the centre of the spherical surface of the microlens and defines the position of the corresponding approximating pinhole.

and the upper limit is:

$$sup(\mathcal{F}_3(k)) = k\phi_r - \frac{\phi_r}{2} + (t-s)\tan\left(\arcsin\left(\frac{1}{n}\sin(\beta_2)\right) + \arcsin\left(\frac{\phi_r}{2r}\right)\right), \tag{3.18}$$

where

$$\beta_2 = \arctan\left(\frac{k\phi_r - \phi_r/2 + ap/2}{a+s}\right) - \arcsin\left(\frac{\phi_r}{2r}\right). \tag{3.19}$$

In these equations, t is the microlens thickness (Figure 3.7), n is the microlens refractive index, $r = F(n-1)$ is the radius of the spherical surface of a microlens, F is the focal distance of a microlens, ϕ_r is the pitch of the microlenses in the recording array, and s is the maximum sag of the microlens:

$$s = r\left(1 - \cos\left(\arcsin\left(\frac{\phi_r}{2r}\right)\right)\right) \tag{3.20}$$

Clearly, according to this model, the energy corresponding to a microlens is spread on the overall area covered by the three fields $\mathcal{F}_1(k)$, $\mathcal{F}_2(k)$, $\mathcal{F}_3(k)$ and the total field is therefore $\mathcal{F}_{refr}(k) = \mathcal{F}_1(k) \cup \mathcal{F}_2(k) \cup \mathcal{F}_3(k)$, whose boundaries are given by:

$$\begin{aligned} inf(\mathcal{F}_{refr}(k)) &= \min_{l=1,2,3}(inf(\mathcal{F}_l(k)))\,, \\ sup(\mathcal{F}_{refr}(k)) &= \max_{l=1,2,3}(sup(\mathcal{F}_l(k)))\,. \end{aligned} \tag{3.21}$$

The highest energy is found at the overlap of the three fields, so on $\mathcal{F}_{inters}(k) =$

$\mathcal{F}_1(k) \cap \mathcal{F}_2(k) \cap \mathcal{F}_3(k)$, limited by:

$$\begin{aligned} inf(\mathcal{F}_{inters}(k)) &= \max_{l=1,2,3}(inf(\mathcal{F}_l(k)))\,, \\ sup(\mathcal{F}_{inters}(k)) &= \min_{l=1,2,3}(sup(\mathcal{F}_l(k)))\,. \end{aligned} \tag{3.22}$$

The imaging fields obtained with the refraction model are compared in the following with the imaging fields obtained with the pinhole approximation, where the pinhole corresponding to each microlens is placed in the centre of the spherical lens. If the distance between the camera aperture and the front of the recording array is denoted by a, then the corresponding pinhole is at a distance $a + r$ in front of the camera, as shown in Figures 3.6-3.7. The width of the pinhole approximation field is comprised between the following limits for each microlens k:

$$\begin{aligned} inf(\mathcal{F}_{p.a.}(k)) &= \frac{ap}{2} + \frac{(c_k - ap/2)(a+t)}{a+r}\,, \\ sup(\mathcal{F}_{p.a.}(k)) &= -\frac{ap}{2} + \frac{(c_k + ap/2)(a+t)}{a+r}\,, \end{aligned} \tag{3.23}$$

where c_k is the y-coordinate of the microlens centre.

The following numerical values for the parameters have been considered:

- camera aperture: ap =180 mm.
- microlens pitch: ϕ_r =1.27 mm.
- microlens refractive index: $n = 1.5$.
- microlens focal distance: $F = 2\phi_r$.
- microlens spherical radius: $r = F(n-1)$.
- microlens thickness: $t = r + f$.
- variable distance between the camera aperture and the recording array: $a \in \{350, 360, 370, 380\}$ ([mm]).

Figure 3.8 presents in parallel the lower and upper limits of the union field $\mathcal{F}_{refr}$, of the intersection field $\mathcal{F}_{inters}$, and of the pinhole approximation field $\mathcal{F}_{p.a.}$ for microlenses whose index k ranges between 0 and 150. A few values of the distance a between the camera aperture and the microlens array have been considered. It

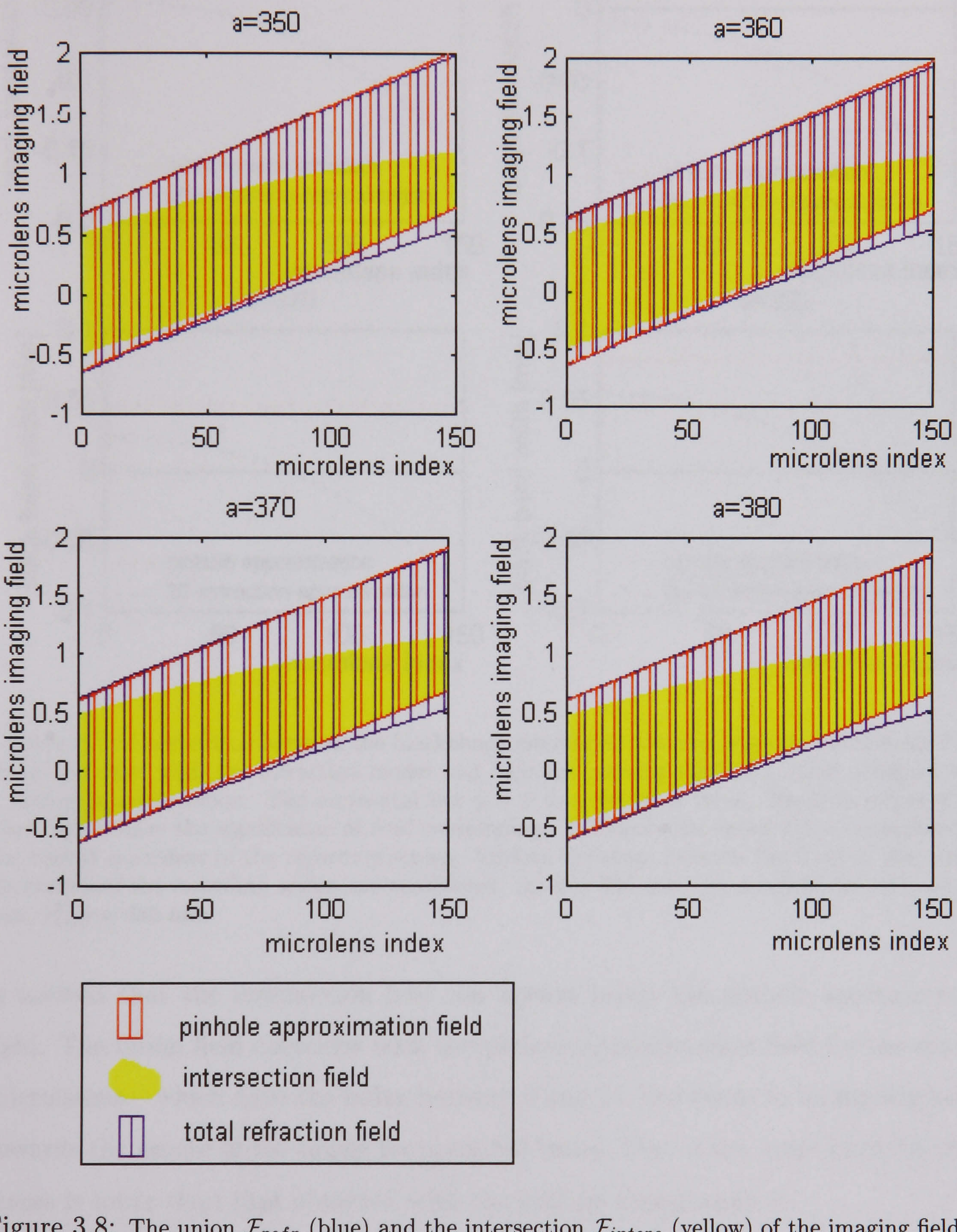


Figure 3.8: The union $\mathcal{F}_{refr}$ (blue) and the intersection $\mathcal{F}_{inters}$ (yellow) of the imaging fields in the refraction model compared with the imaging field $\mathcal{F}_{p.a.}$ (red) obtained with a pinhole approximation. 0 on the ordinate axis represents the centre of the microlens. The variation of the width of the three imaging fields versus the microlens index; the microlens index $k = 0$ corresponds to the central microlens of the recording screen. Various distances between the front of the camera aperture and the recording screen are considered: (a) $a = 350$ mm, (b) $a = 360$ mm, (c) $a = 370$ mm, (d) $a = 380$ mm.

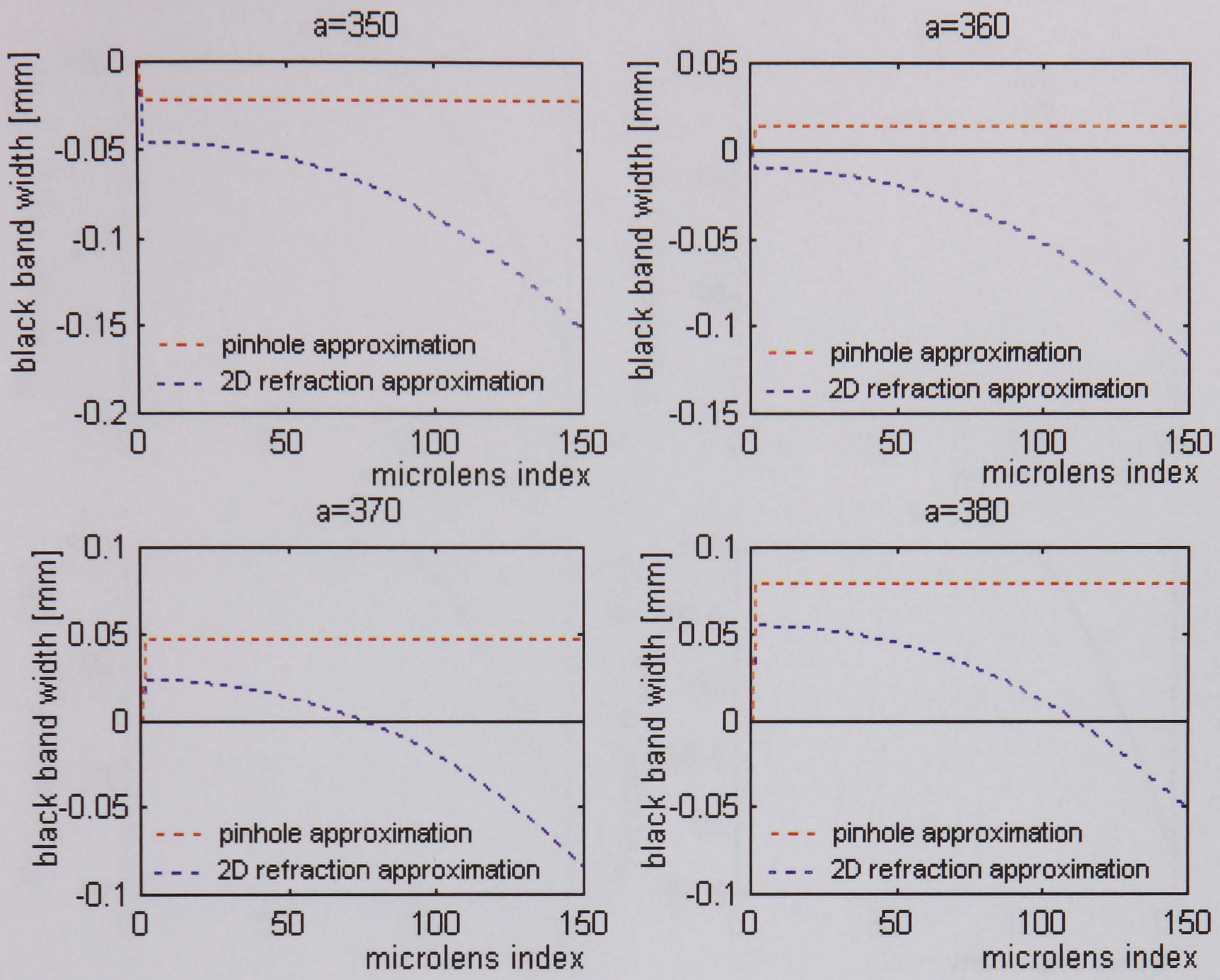


Figure 3.9: Comparison between the blackbands interleaved between adjacent union fields $\mathcal{F}_{refr}$ (blue) obtained with the refraction model and adjacent imaging field $\mathcal{F}_{p.a.}$ (red) obtained with a pinhole approximation. The horizontal line $y = 0$ is a reference value. Negative values of the black bands have the significance of field overlapping. The microlens index $k = 0$ corresponds to the central microlens of the recording screen. Various distances between the front of the camera aperture and the recording screen are considered: (a) $a = 350$ mm, (b) $a = 360$ mm, (c) $a = 370$ mm, (d) $a = 380$ mm.

is noticed that the intersection field lies always inside the pinhole approximation field. The union field coincides with the pinhole approximation field for the central microlenses – which have the index between 0 and 75, but tends to be slightly larger towards the centre of the image for marginal lenses, that is the lower limit for these lenses is lower than that obtained with the pinhole approximation.

Figure 3.9 shows in parallel the width of the black bands between subimages corresponding to adjacent microlenses for the refraction model:

$$w_{bb}^{refr}(k) = inf(\mathcal{F}_{refr}(k)) - sup(\mathcal{F}_{refr}(k-1))\,, \tag{3.24}$$

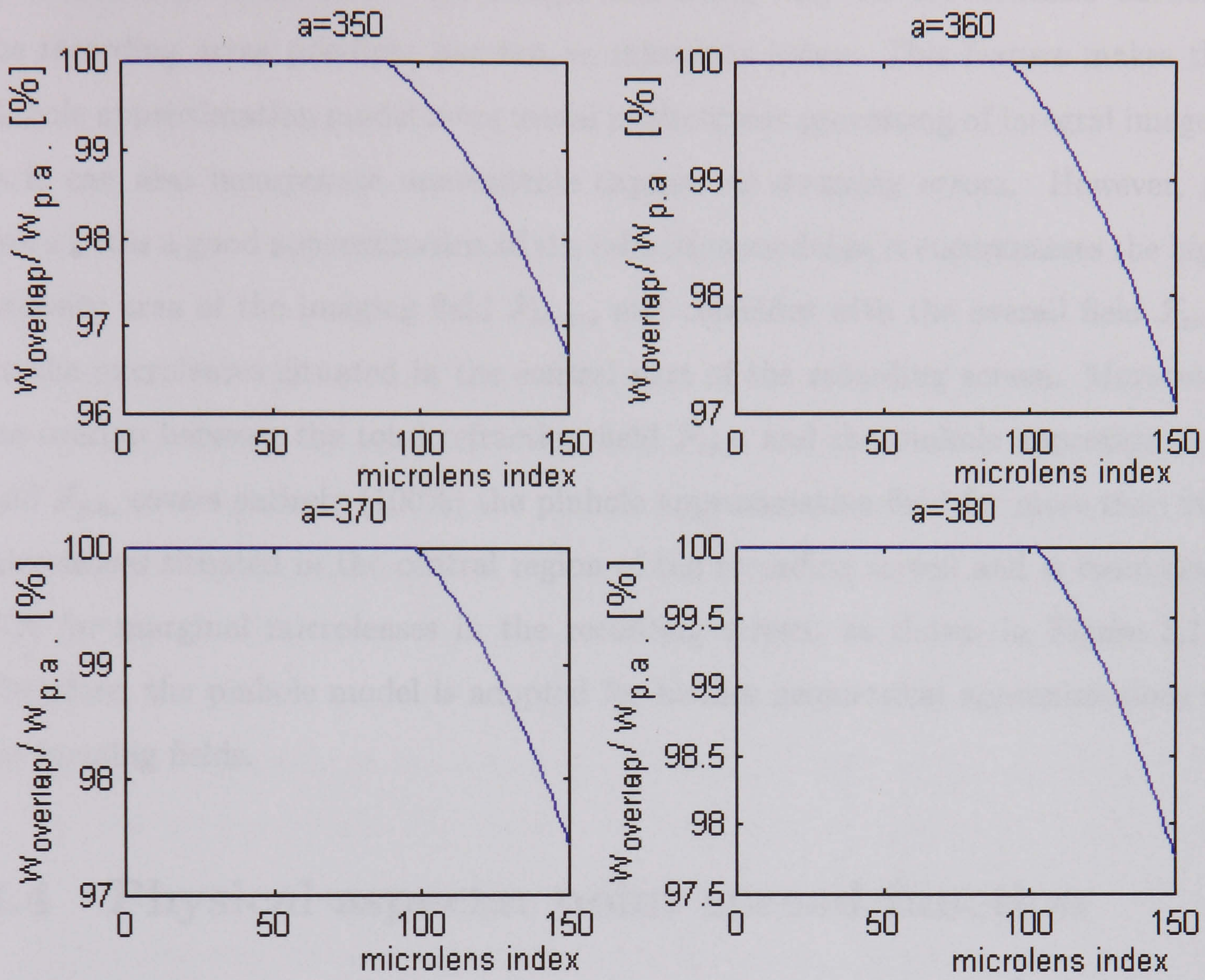


Figure 3.10: Ratio in percentages between the width of the overlap $\mathcal{F}_{refr} \cap \mathcal{F}_{p.a.}$ and the width of the field obtained with a pinhole approximation $\mathcal{F}_{p.a.}$. The microlens index $k = 0$ corresponds to the central microlens of the recording screen. Various distances between the front of the camera aperture and the recording screen are considered: (a) $a = 350$ mm, (b) $a = 360$ mm, (c) $a = 370$ mm, (d) $a = 380$ mm.

and for the pinhole approximation:

$$\begin{aligned} w_{bb}^{p.a.}(k) &= inf(\mathcal{F}_{p.a.}(k)) - sup(\mathcal{F}_{p.a.}(k-1)) \\ &= \frac{ap(t+r)}{a+r} = const\,. \end{aligned} \tag{3.25}$$

The negative values of the width of the black bands mean that the adjacent imaging fields actually overlap. It is noticed that the refraction imaging fields tend to overlap for marginal microlenses. At the same time, the farthest the recording plane from the camera aperture is, the lower the overlap. However, in the refraction model, both the imaging fields and the black bands between them have widths which depend on the microlens index and on the distance a. This differs from the pinhole approximation model, where the imaging fields and the black bands

have constant width across the image, depending only on the distance between the recording array position, but not on microlens index. This feature makes the pinhole approximation model more useful in electronic processing of integral images, as it can also incorporate unavoidable capture or scanning errors. However, as proved, it is a good approximation of the refraction model as it encompasses the high intensity area of the imaging field $\mathcal{F}_{inters}$ and coincides with the overall field $\mathcal{F}_{refr}$ for the microlenses situated in the central part of the recording screen. Moreover, the overlap between the total refraction field $\mathcal{F}_{refr}$ and the pinhole approximation field $\mathcal{F}_{p.a.}$ covers entirely (100%) the pinhole approximation field for more than 200 microlenses situated in the central region of the recording screen and is more than 97% for marginal microlenses in the recording screen, as shown in Figure 3.10. Therefore, the pinhole model is adopted for further geometrical approximations of the imaging fields.

3.4 Physical aspects: point spread function

3.4.1 Optical imaging model

The model of image formation considered was reported by Rayleigh ([80]) and enhanced by Duffieux ([27]), and is applicable to both coherent and incoherent illumination conditions. It visualises an image as the combination of the Airy patterns (or more complicated patterns if aberrations are involved) that an optical system would separately produce with the light leaving every individual point in the object ([80], [27], [85]). In Rayleigh's model, each point being imaged is considered a light source itself. The light rays emerging from it are limited in their entry into the imaging system by the finite aperture of the imaging lens, and the diffraction pattern of that aperture is formed in the plane conjugate to the plane containing the point source, i.e. in the image plane of the object points as light sources. Each point is therefore imaged not as a point, but as the Airy pattern of the aperture of the imaging lens (or more complicated patterns if aberrations are present) ([85]). The centre of the Airy pattern of a point $Q(x_0, y_0)$ can be determined and is, as expected, its conjugate according to the geometrical optics, that is the point $Q'(mx_0, my_0)$, where m is the

transverse magnification ([3], [45]) .

If the illumination is incoherent, the Airy patterns due to all the object points are simply additive. By considering an object in the form of a row of equally spaced point sources, Rayleigh used a Fourier series to represent the summation of the regularly spaced Airy patterns forming the image according to this model ([80], [85]).

Duffieux ([27]) made an important development of the model by formally expressing the image of a continuous object distribution as a convolution of that distribution with the ***point spread function (PSF)***, or impulse response of the system (the Airy pattern in the simplest case), using complex amplitudes and intensities as appropriate for the illumination conditions. Application of the convolution theorem demonstrates that the spatial frequency spectrum of the image is the product of the frequency spectrum of the object distribution and that of the response of the system. The optical system can therefore be considered as transferring, to different extents, the spatial frequency components of the object to the image plane ([85]).

Due to the very fine pitch of a microlens, diffraction plays a major role during the optical process that leads to the formation of the integral image. Therefore, the Rayleigh-Duffieux model of image formation, based on the diffraction phenomenon, is reckoned to be the most appropriate for describing the optical process involved in the 3D-integral recording system ([56], [57]). The objects imaged through the integral system are illuminated under normal conditions, so the illumination involved is essentially incoherent. Therefore all mutual coherence between light from two separate points in the object is neglected, and the flux density is obtained by summing the individual contributions from all points in the object ([56], [57]).

3.4.2 Point spread function in the 3D-lenticular integral recording system

As a result of the optical process, each point will give rise to an intensity distribution in the image: this is *the point spread function.* It will be the same for all object points in a sufficiently small neighbourhood. The point spread function is defined ([3]) as the square of the modulus of the complex amplitude $\mathcal{A}$ of the light in a

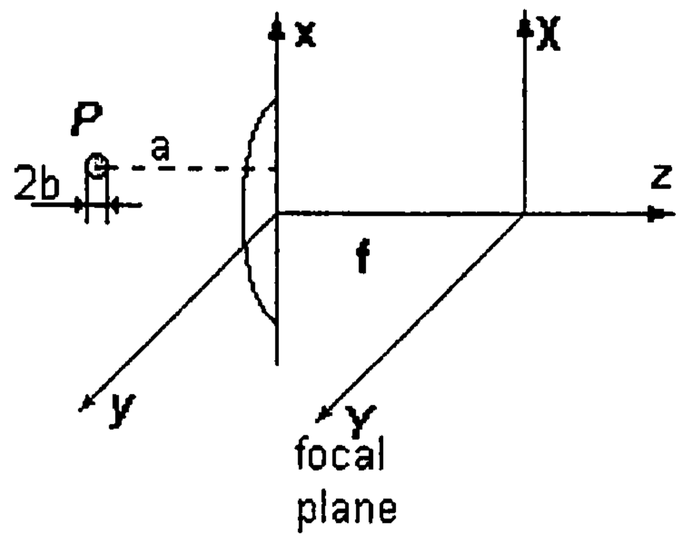


Figure 3.11: Close and remote imaging.

reference plane spanned by an XY Cartesian frame:

$$K(X,Y) = |\mathcal{A}(X,Y)|^2 . \tag{3.26}$$

The calculation of the point spread function of the lenticular integral imaging system has to take into account the three microlens arrays through which light passes prior to the integral reconstruction and recording of the imaged object.

The first mathematical treatment of a single tier integral imaging system was presented by Burckhardt et al. in the context of both *remote* ([13]) and *close imaging* ([14]). The calculations were performed for a square based microlens, in a direction parallel to one side of the square lenslet. The present approach considers square based spherical surface microlens arrays in the ATS and cylindrical surface microlenses in the recording screen, all calculations being performed two dimensionally on the microlens aperture. In order to derive the point spread function, a quasi-pointwise source of light P situated at a distance a in front of a microlens is considered (see Figure 3.11). The half-"thickness" of this light source is denoted by b. The image formation is affected by so called *close imaging* when the microlens screen is located in the immediate vicinity of the object point (such that a and b are comparable), whilst *remote imaging* operates when the distance of the imaged object to the microlens screen is relatively large compared to the object thickness ($a \gg b$). The integral image formation and recording of deep scenes combines both remote and close imaging, therefore the analysis of each is necessary.

Using the Fresnel - Kirchhoff equation, Burckhardt et al. derived the equation

of the point spread function for a microlens screen along the lateral axis Ox for both remote ([13]) and close imaging conditions ([14]). Adapting this result to the lenticular integral camera system, which uses square based lenslets in the ATS, the point spread function of the first microlens array in the focal plane is ([56]):

$$K_1(X,Y) = \exp\left(-\frac{X^2+Y^2}{u^2}\right) \tag{3.27}$$

where the coefficient u is given under remote imaging conditions by:

$$u_{rm}^2 = 0.104\frac{\phi^2 b^2 f^2}{a^4} + 0.245\frac{\lambda^2 f^2}{\phi^2}\,, \tag{3.28}$$

and for close imaging conditions by:

$$u_{cl}^2 = 0.104\frac{\phi^2 f^2}{a^2} + 0.245\frac{\lambda^2 f^2}{\phi^2}\,. \tag{3.29}$$

The spread factor u^2 depends on camera parameters and the position of a point in space ([56]). In the expression, ϕ is the pitch, f is the focal distance of a microlens, b is the thickness of a relevant portion of scene which contains point P, and a is the distance between an intermediary image and a microlens array, given by:

$$a = \frac{-z_P f_1 + z_P S + S^2}{f_1 + z_P + S} - f\,.$$

The thickness parameter b can be chosen in accordance with the depth resolution of the image. In all simulations presented in this paper, b has been assumed to be equal to 1 mm. Analysing u^2, it is noticed that the first term of the sum expresses the defocus influence, whilst the second one is due to diffraction ([74]). Due to the symmetry of the optical process that occurs at the central double microlens array, the point spread function of the second microlens array is also given by (3.27). Therefore, the spread function of the spatial inversion process (associated with the central screen) is given by the convolution of (3.27) with itself ([56]):

$$K_{ATS} = K_1 \star K_1\,, \tag{3.30}$$

which leads to:

$$\begin{aligned} K_{ATS}(X,Y) &= \iint_{R^2} \exp\left(-\frac{X'^2}{u^2}\right) \exp\left(-\frac{Y'^2}{u^2}\right) \\ &\quad \times \exp\left(-\frac{(X-X')^2}{u^2}\right) \exp\left(-\frac{(Y-Y')^2}{u^2}\right) dX'\,dY' \\ &= \int_{-\infty}^{\infty} \exp\left(-\frac{X'^2+(X-X')^2}{u^2}\right) dX' \\ &\quad \times \int_{-\infty}^{\infty} \exp\left(-\frac{Y'^2+(Y-Y')^2}{u^2}\right) dY' \,. \end{aligned} \tag{3.31}$$

The two integrals can be computed as follows:

$$\begin{aligned} \int_{-\infty}^{\infty} \exp\left(-\frac{X'^2+(X-X')^2}{u^2}\right) dX' &= \exp\left(-\frac{X^2}{2u^2}\right) \\ &\quad \times \int_{-\infty}^{\infty} \exp\left(-\frac{2(X'-(1/2)X)^2}{u^2}\right) dX' \\ &= \exp\left(-\frac{X^2}{2u^2}\right) \frac{u}{\sqrt{2}} \int_{-\infty}^{\infty} \exp(-t^2)\,dt \\ &= \frac{\pi \cdot u}{2\sqrt{2}} \exp\left(-\frac{X^2}{2u^2}\right) . \end{aligned} \tag{3.32}$$

Computations yield for K_{ATS} the following Gaussian distribution ([56]):

$$K_{ATS}(X,Y) = Cv^2 \exp\left(-\frac{X^2+Y^2}{v^2}\right) , \tag{3.33}$$

where C is a constant which is omitted in further computations, and:

$$v^2 = 2u^2 \,. \tag{3.34}$$

The point spread function associated with the recording cylindrical microlens screen in the focal plane (which is the plane of the photographic plate) has an expression analogous to (3.27), but contains different parameters:

$$K_{rec}(X,Y) = \exp\left(-\frac{X^2}{w_1^2}\right) \exp\left(-\frac{Y^2}{w_2^2}\right) , \tag{3.35}$$

where:

$$w_1^2 = w_{1,rm}^2 = 0.104\frac{\psi_r^2 b^2 F^2}{a_r^4} + 0.245\frac{\lambda^2 F^2}{\psi_r^2}\,, \qquad (3.36)$$

$$w_2^2 = w_{2,rm}^2 = 0.104\frac{\phi_r^2 b^2 F^2}{a_r^4} + 0.245\frac{\lambda^2 F^2}{\phi_r^2}\,,$$

for remote imaging conditions, and:

$$w_1^2 = w_{1,cl}^2 = 0.104\frac{\psi_r^2 F^2}{a_r^2} + 0.245\frac{\lambda^2 F^2}{\psi_r^2}\,, \qquad (3.37)$$

$$w_2^2 = w_{2,cl}^2 = 0.104\frac{\phi_r^2 F^2}{a_r^2} + 0.245\frac{\lambda^2 F^2}{\phi_r^2}\,,$$

for close imaging conditions.

In the above formulae, ψ_r, ϕ_r, and F are the length, width, and the focal distance of a recording microlens, whilst a_r is the distance between the reconstructed model P' and the recording microlens array ($a_r = D - |z_P|$), and λ is the wavelength.

The spread function of the entire optical process combines the spread effect of both the central double microlens screen and the recording array. Therefore, it is the convolution of K_{ATS} and K_{rec}:

$$K_{total} = K_{ATS} \star K_{rec}\,, \qquad (3.38)$$

which gives ([56]):

$$K_{total}(X,Y) = v^2 \iint_{R^2} \exp\left(-\frac{X'^2}{v^2} - \frac{(X-X')^2}{w_1^2}\right) \exp\left(-\frac{Y'^2}{v^2} - \frac{(Y-Y')^2}{w_2^2}\right) dX'\,dY'\,. \qquad (3.39)$$

Computations yield ([56]):

$$K_{total}(X,Y) = \alpha(z)\exp\left(-\frac{X^2}{a_1^2}\right)\exp\left(-\frac{Y^2}{a_2^2}\right)\,, \qquad (3.40)$$

where

$$\alpha(z) = \frac{v^4 w_1^2 w_2^2}{\sqrt{(v^2+w_1^2)(v^2+w_2^2)}}\,,$$

$$a_1^2 = v^2 + w_1^2\,, \qquad a_2^2 = v^2 + w_2^2\,.$$

Equation (3.40) expresses the spread function in local coordinates associated with a microlens. Using the global Cartesian system depicted in Figure 3.1, the point spread behind a specific microlens k of the recording array, imaging a point P can be written as ([56]):

$$K_{total,k}(x,y) = \alpha(z)\exp\left(-\frac{(x-x'_k)^2}{a_1^2}\right)\exp\left(-\frac{(y-y'_k)^2}{a_2^2}\right) , \tag{3.41}$$

where x'_k, and y'_k are the coordinates of the recorded intensity distribution P'_k behind microlens k, given in the previous section.

The overall point spread function is the sum of the point spreads behind all the microlenses which 'see' the physical point P. Therefore ([56]):

$$K_{global}(x,y) = \sum_k K_{total,k}(x,y) , \tag{3.42}$$

where the summation is done with respect to the indices k of the microlenses involved in recording the point P.

3.5 Depth information encoded in the unidirectional integral image

3.5.1 Number of recorded intensity distributions corresponding to a physical point

An important consideration is the number of recorded intensity distributions of the optical reconstruction P' of a physical point P (Figure 3.4(b)). Computations show that this number depends on both lateral position and depth of the point.

For a point P whose reconstruction P' is situated in front of the recording screen, the light beams emerging from P' are spread laterally on the y-interval delimited by:

$$\max\left(y_{scr}, ap + (y_P - ap)\frac{D-S}{|z_P|-S}\right) \tag{3.43}$$

and

$$\min\left(y_{scr} + m\cdot\phi_r, y_P\frac{D-S}{|z_P|-S}\right) , \tag{3.44}$$

where *ap* denotes the camera aperture, m is the number of microlenses in the recording screen, and y_{scr} describes the lateral position of the recording screen (Figure 3.4(b)). Analogously, the light beams which contribute to the reconstruction of a point behind the recording screen cover the interval given by:

$$\max\left(y_{scr}, ap + y_P\frac{D-S}{|z_P|-S}\right), \tag{3.45}$$

and

$$\min\left(y_{scr} + m\cdot\phi_r, ap + (y_P - ap)\frac{D-S}{|z_P|-S}\right), \tag{3.46}$$

These intervals depend both on the lateral (y_P) and depth (z_P) position of the point, and also on the position of the optical reconstruction relative to the recording screen ($D - |z_P|$).

The maximum recorded image number for a given depth is attained for points having a central position in the imaged scene:

$$N_{max} = \left[ap\frac{D-|z_P|}{\phi_r(|z_P|-S)}\right] + 1\,. \tag{3.47}$$

This number depends on the camera aperture *ap*. Marginal points have obviously fewer recorded images due to the finite dimensions of the recording screen. A Matlab application has been used to plot the step functions that describe the dependence between the number of recorded intensity distributions of a point and depth for different lateral positions of the point (Figure 3.12). Figure 3.12 depicts the number of recorded intensity distributions for marginal points (a), respectively central points (b), with depth ranging between 0.4 m and 0.58 m when the recording screen is at a distance D=0.58 m from the central plane of a camera having an aperture *ap*=0.3 m. As can be seen, points in the proximity of the recording plane, at a distance smaller than the focal distance of a recording microlens, are not recorded. This singularity is explained in terms of physical optics in section 3.5.2.

3.5.2 The influence of the point depth on the point spread

The shape of K_{total}, which is a two variable Gaussian function (Figure 3.13), suggests that the energy forming the recorded intensity distribution is concentrated on a

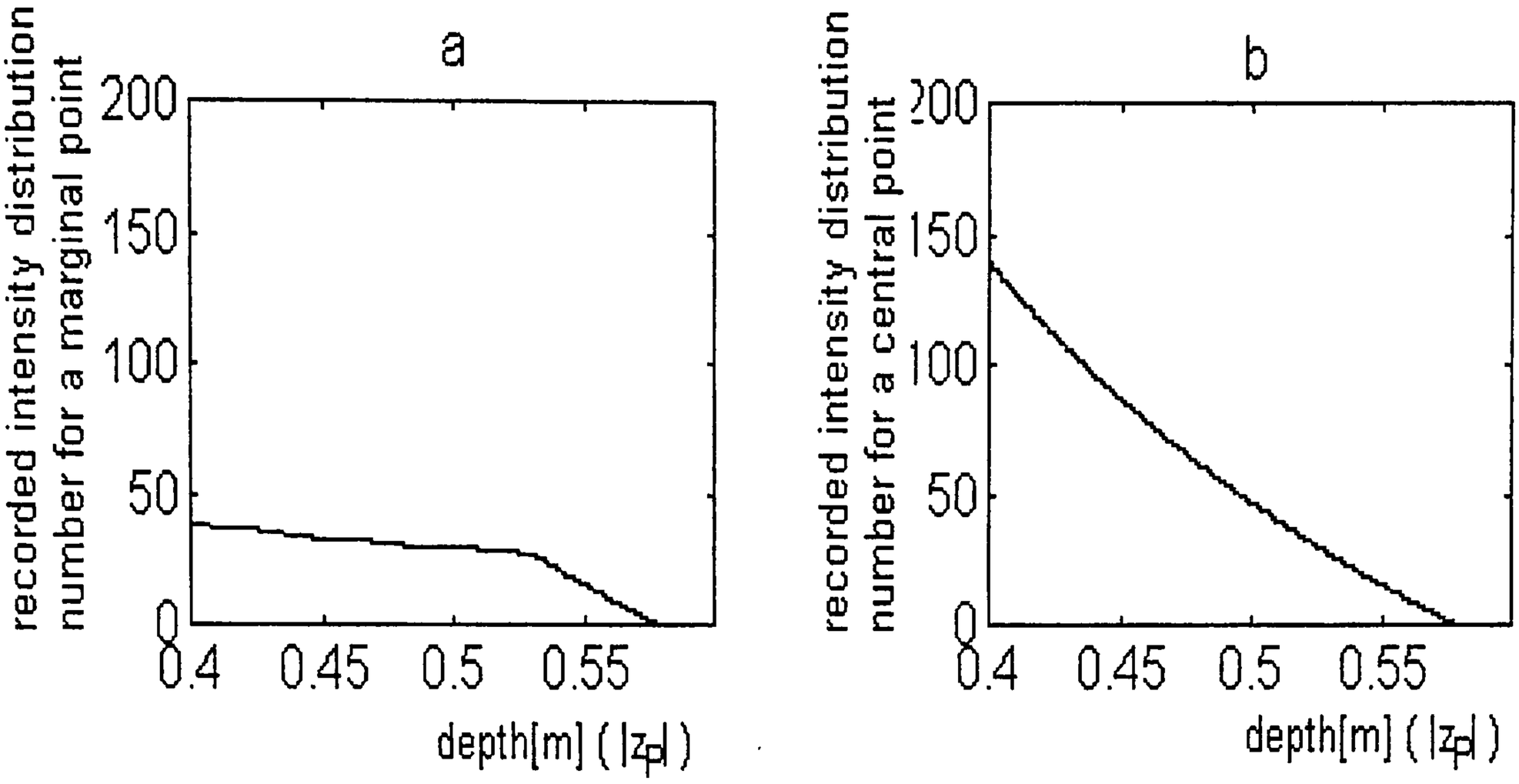


Figure 3.12: The number of recorded intensity distributions of a point with respect to its depth for a marginal (a), and a central point (b).

rectangular spot whose dimensions depend on point depth. Matlab applications have been used to plot the dependence on depth for both the width and the length of a rectangle that concentrates more than 98% of the intensity. Figure 3.14 presents results corresponding to a recording screen position D =0.58 m. The wavelength considered in the computations is $\lambda = 500$ nm. The graphs prove that the dimensions attain a stability level for reconstructed points situated at more than 3 cm away from the recording screen, and that the spread is larger when the recorded points are nearer the recording screen. For the points reconstructed in the immediate vicinity of the recording plane (at a distance smaller than 1.5 mm) , the spread is infinite, therefore these points do not contribute stricto sensu to the recorded spatial information. This result confirms that the number of contributing intensity distributions for these points is 0, as proved on a geometrical basis in section 3.5.1.

3.6 Conclusions

The operation of the 3D-unidirectional (lenticular) camera system, which uses cylindrical surface microlenses to produce images with continuous horizontal parallax, and the associated image formation and recording processes have been presented. A mathematical model of the optical process involved by the 3D-unidirectional (lenticular) integral recording camera system, comprising both geometrical and physical

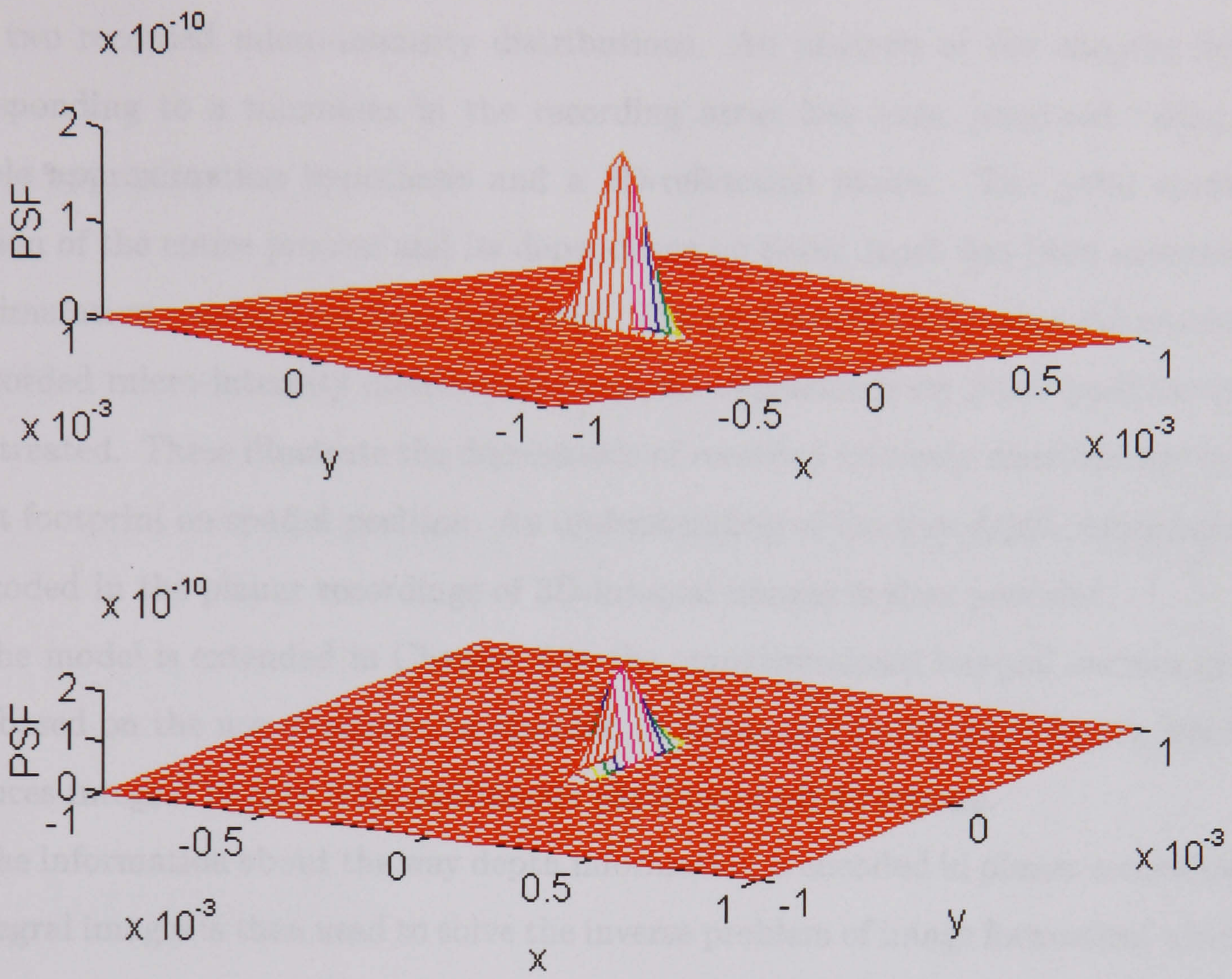


Figure 3.13: The shape of the point spread function behind a cylindrical surface recording microlens.

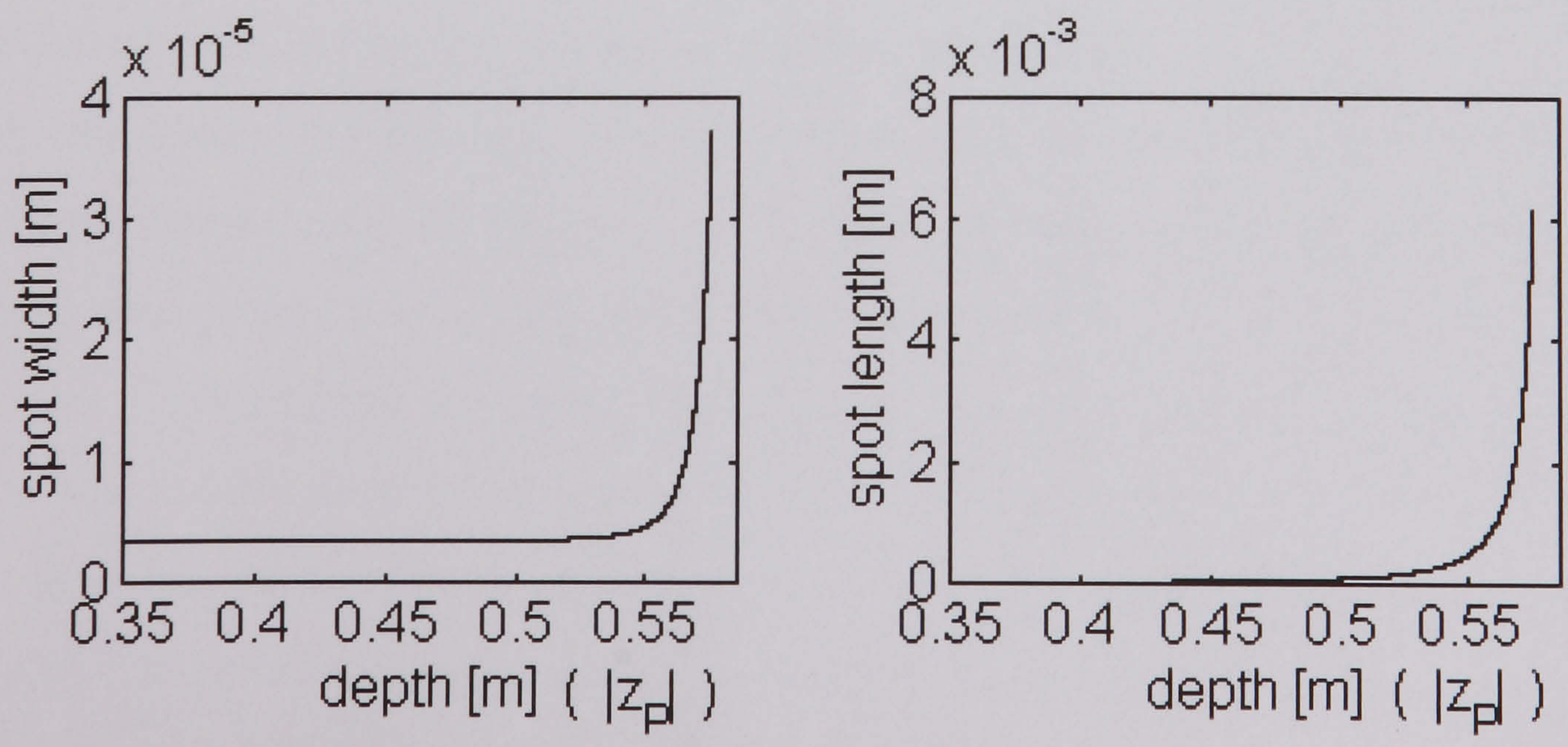


Figure 3.14: The dependence on depth for the width and length of a point's recorded image.

aspects, has been described. The coordinates of the intermediary, integrated, and recorded data have been computed and used to retrieve an object point position from two recorded micro-intensity distributions. An analysis of the imaging field corresponding to a microlens in the recording array has been proposed, using a pinhole approximation hypothesis and a 2D-refraction model. The point spread function of the entire process and its dependence on point depth has been analysed, and simulation results have been presented. The problem of computing the number of recorded micro-intensity distributions and the dependence on point position has been treated. These illustrate the dependence of recorded intensity distributions and object footprint on spatial position. An understanding of the way depth information is encoded in the planar recordings of 3D-integral images is thus provided.

The model is extended in Chapter 4 to the omnidirectional integral camera system, based on the use of square, or circular, or hexagonal based microlenses, which produces integral images with continuous parallax in all directions.

The information about the way depth information is encoded in planar recordings of integral images is then used to solve the inverse problem of image formation, which is to extract the composition of the recorded scene and draw the depth map of the object space (Chapters 5-7).

Chapter 4

Mathematical Model of the Omnidirectional Integral Imaging System

4.1 Introduction

The two tier camera system can be designed such that the resultant 3D-integral image has continuous parallax in all directions, not only horizontally, as it is the case of the lenticular integral system discussed in Chapter 3. The applications of this configuration are multiple, ranging from 3D-television to robotic vision and medical displays, so that the discovery of the rules relating the system parameters to the formation and quality of the 3D-image is a crucial problem.

For this reason, a mathematical model of the omnidirectional 3D-integral recording camera system using either square, or circular, or hexagonal based spherical surface microlens arrays in the autocollimating transmission screen and encoding screen is derived. This analysis is a generalisation of the work reported in Chapter 3 and [56] for the unidirectional (lenticular) system. The close relationship between the image morphology and the three dimensional recorded scene composition is analysed, and comparative considerations for the three variants of the system are made. The relationship between the recorded intensity distributions and the point depth is particularly emphasised.

To the best of our knowledge, to date, this is the only reported ([57]) comprehensive analysis of an integral imaging system in terms of depth influence on point spread.

The present work considers three variants of the system, depending on the type of microlenses used in the ATS and recording array:

- square based microlenses
- circular based microlenses
- hexagonal based microlenses

Two types of microlens array packaging are considered:

- rectangular net – for square and circular based microlenses
- honeycomb – for circular and hexagonal based microlenses

4.2 Geometrical aspects

The configuration of the camera system resembles very well to the one used in the previous chapter, the salient difference being that the recording array uses a two dimensionally active configuration employing either square, or circular, or hexagonal based microlenses, instead of cylindrical microlenses. However, as in the case of unidirectional (lenticular) integral imaging, all the equations in this chapter are related to the Cartesian coordinate system $Oxyz$ depicted in figure 4.1, where the z-axis denotes the depth direction, and the x, y-axes describe the lateral positions. The macro- and microlenses are indexed in the sense indicated by the positive axis Oy.

A physical point $P(x_p, y_p, z_p)$ imaged by the integral camera system has a number of states prior to the recording stage. The macrolenses of the first array present a first set of images of P near the autocollimating transmission screen. These images

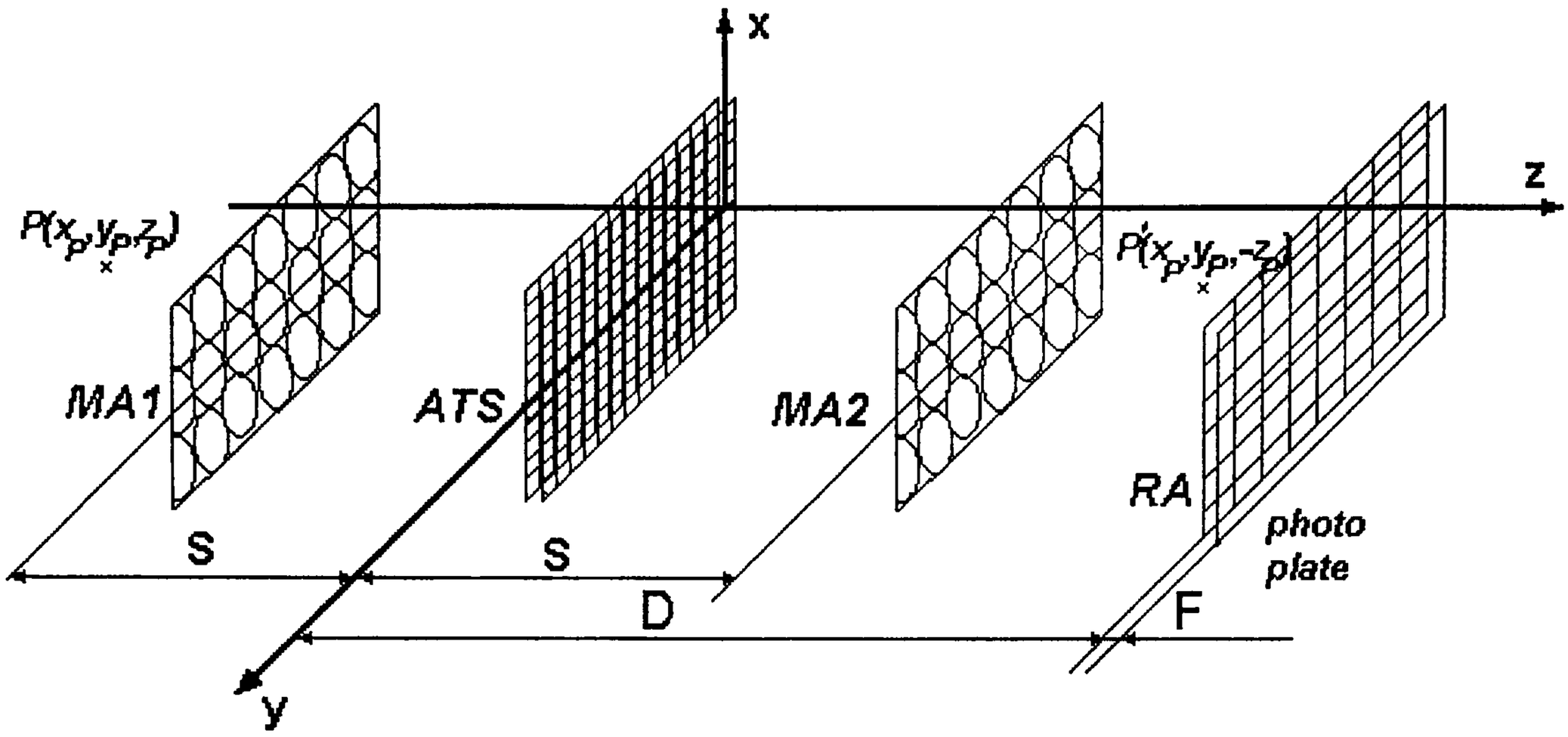


Figure 4.1: The two tier optical system

are located at the points P_k^1 whose coordinates are ([57]):

$$\begin{aligned} x_k^1 &= \frac{a_k(z_P + S) + x_P f_1}{f_1 + S + z_P} \\ y_k^1 &= \frac{b_k(z_P + S) + y_P f_1}{f_1 + S + z_P} \\ z_k^1 &= \frac{z_P f_1 - z_P S - S^2}{f_1 + S + z_P}, \end{aligned} \tag{4.1}$$

for any macrolens k in the first array, where: f_1 is the focal length of the macrolenses, (a_k, b_k, S) denote the coordinates of macrolens k 's centre, and S is the distance between a macrolens array and the central plane Oxy (see Figure 4.1 and Figure 4.2(a)).

The intermediary images P_k^1 become object points for the ATS. The light emerging from them is refracted and diffracted through the microlens apertures of the first array, producing an intensity distribution in the common focal plane of the two microlens arrays. The area between microlenses is assumed opaque. This intensity distribution is then transmitted by the second microlens array through a symmetric optical process. As a result, the ATS operates by refracting the transmitted ray at an equal angle of opposite sign to the angle of incidence of the impinging ray. In this way, every "object" point P_k^1 in front of the screen has a quasi-equidistant voxel

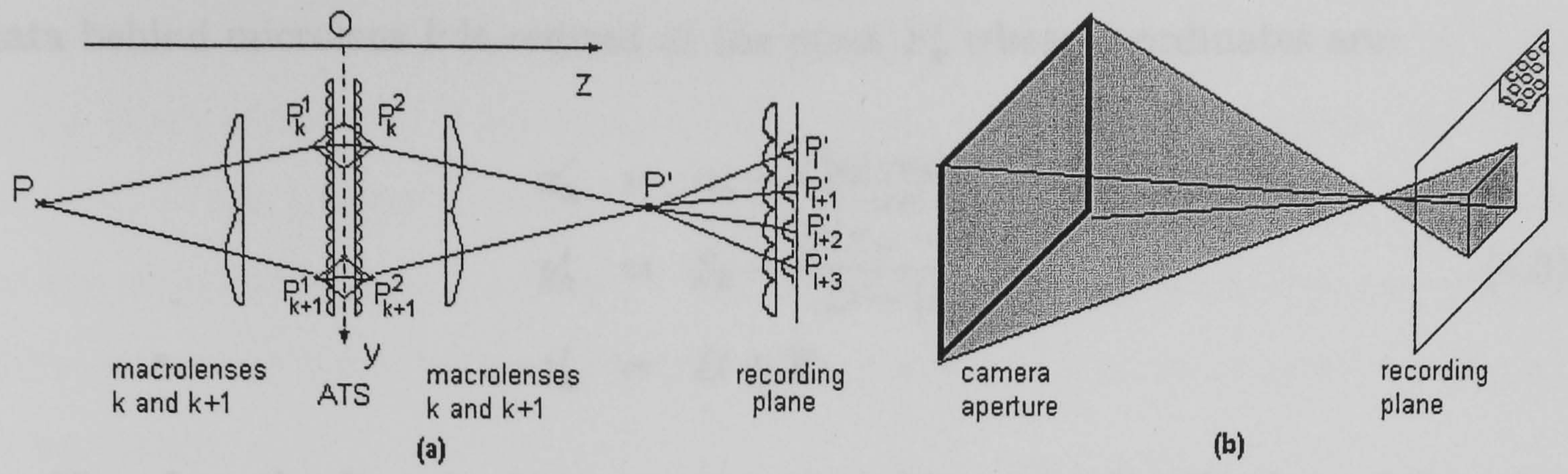


Figure 4.2: Transverse sections of the camera arrangement with a plane $x = x_P$: (a) Schematic representation of the intermediary, integral and recorded optical models of a physical point P; (b) Microlenses of the recording array that record a certain optical reconstruction P'.

representation centred at the point $P_k^2(x_k^2, y_k^2, z_k^2)$ behind the screen, where ([57]):

$$\begin{aligned} x_k^2 &= \frac{a_k(z_P + S) + x_P f_1}{f_1 + S + z_P} \\ y_k^2 &= \frac{b_k(z_P + S) + y_P f_1}{f_1 + S + z_P} \\ z_k^2 &= \frac{-z_P f_1 + z_P S + S^2}{f_1 + S + z_P} \end{aligned} \tag{4.2}$$

Any of the intermediary voxels P_k^2 is then collected by the corresponding output macrolens k and imaged at the location $P'(x_P, y_P, -z_P)$ in space, which is the equal conjugate image location of point P with respect to the Oxy plane. It does not therefore depend upon the macrolens k, and the result is an integration of directional information about P which is the result of the optical process through all the macro- and microlenses that have imaged it. This creates the parallax in all directions visible in the final integrated optical model P' ([57]).

The intensity distributions related to the integral image P' of P may be recorded on a photographic plate that lies behind the recording microlens array (containing microlenses of focal distance F, situated at a distance D from the central Oxy plane - Figure 4.1). P' has recorded intensity distributions in a number of microlenses, the number being determined by the axial position of P'.

If a microlens k, of centre $C_k(\alpha_k, \beta_k, D)$, "sees" the point P', then the recorded

data behind microlens k is centred at the point P'_k whose coordinates are:

$$\begin{aligned} x'_k &= \alpha_k - \tfrac{F(x_P-\alpha_k)}{D-|z_P|} \\ y'_k &= \beta_k - \frac{F(x_P-\beta_k)}{D-|z_P|} \\ z'_k &= D+F\,. \end{aligned} \tag{4.3}$$

Therefore, the disparity between two recorded intensity distributions of the point P corresponding to microlenses k and j, of centres $C_k(\alpha_k,\beta_k,D)$ and $C_j(\alpha_j,\beta_j,D)$, can be computed as:

$$d_{jk} = \left|\frac{D+F-|z_P|}{D-|z_P|}\right| \delta_{jk}\,, \tag{4.4}$$

where $\delta_{jk} = \sqrt{(\alpha_j-\alpha_k)^2+(\beta_j-\beta_k)^2}$ is the distance between C_j and C_k.

It allows the recovery of the physical point's position when two of its intensity distributions, centred at $P'_j(x'_j,y'_j,D+F)$ and $P'_k(x'_k,y'_k,D+F)$, are known (and, therefore, the disparity d_{jk} between them is known). Namely ([57]):

$$\begin{aligned} x_P &= \tfrac{\alpha_k d_{jk}-x'_k\delta_{jk}}{d_{jk}-\delta_{jk}} \\ y_P &= \frac{\beta_k d_{jk}-x'_k\delta_{jk}}{d_{jk}-\delta_{jk}} \\ z_P &= \frac{F\delta_{jk}}{d_{jk}-\delta_{jk}} - D\,. \end{aligned} \tag{4.5}$$

4.3 Physical aspects: point spread function

Generalising the mathematical treatment of the unidirectional (lenticular) two tier integral camera system presented in [56] and Chapter 3, a novel analytical model of the omnidirectional integral camera system has been developed and reported ([57]). The present approach considers square, circular, and hexagonal based spherical microlens arrays, all calculations being performed two dimensionally on the microlens aperture. In order to derive the point spread function, a quasi-pointwise source of light P situated at a distance a in front of a microlens is considered (see Figure 3.11). The half-"thickness" of this light source is denoted by b.

As in the unidirectional camera case, both close and remote imaging conditions are considered. We remind the reader that the image formation is affected by *close*

imaging when the microlens screen is located in the immediate vicinity of the object point (such that a and b are comparable), whilst *remote imaging* occurs when the distance of the imaged object to the microlens screen is relatively large compared to the object thickness ($a \gg b$). The integral image formation and recording of deep scenes combines both remote and close imaging, therefore the analysis of each is necessary.

4.3.1 Point spread function for square based microlens integral system

Close imaging and square based microlenses

First, square based spherical surface microlenses of pitches ϕ and respectively ϕ_r are considered in the ATS and in the recording microlens array. Burckhardt et al. ([14]) derived the formula of the point spread function for a microlens screen along the lateral axis Ox at close imaging from the Fresnel-Kirchhoff integral equation. Adapting this result to the integral camera system, which uses square based lenslets in the ATS, the point spread function of the first microlens array in the focal plane is:

$$K_{sq,cl}(X,Y) = C\alpha_{sq,cl} \exp\left(-\frac{X^2+Y^2}{u^2_{sq,cl}}\right) \tag{4.6}$$

where C denotes a constant and:

$$\alpha_{sq,cl} = \frac{a^2}{4.16\phi^4 f^4 \lambda^2 + \pi^2 b^2}\,, \tag{4.7}$$

and the coefficient:

$$u^2_{sq,cl} = 0.104\frac{\phi^2 f^2}{a^2} + 0.245\frac{\lambda^2 f^2}{\phi^2}\,, \tag{4.8}$$

depend on camera parameters and the position of a point in space. In the expression, ϕ is the microlens pitch, f is the focal distance of a microlens, and a is the distance between an intermediary image and a microlens plane, given by:

$$a = a_{ATS} = \frac{-z_P f_1 + z_P S + S^2}{f_1 + z_P + S} - f\,.$$

It can be noticed that the first term of the sum in $u^2_{sq,cl}$ expresses the defocus influence, whilst the second one is due to diffraction ([74]). This is a situation analogous to that discussed in Chapter 3 and [56] for cylindrical surface microlenses. It is of crucial importance for the present research to notice that both $u_{sq,cl}$ and $\alpha_{sq,cl}$ contain depth information about the point depth.

The point spread function of the second microlens array is also given by (4.6) because the optical process that occurs at the central double microlens array is symmetric. Therefore, the spread function of the spatial inversion process (associated to the central screen) is given by the convolution of (4.6) with itself:

$$K^{ATS}_{sq,cl} = K_{sq,cl} \star K_{sq,cl} \,.$$

Straightforward computations yield for $K^{ATS}_{sq,cl}$ the next Gaussian distribution:

$$K^{ATS}_{sq,cl}(X,Y) = C\left(\alpha^{ATS}_{sq,cl}\right)^2 u^{ATS}_{sq,cl} \exp\left(-\frac{X^2+Y^2}{2\left(u^{ATS}_{sq,cl}\right)^2}\right), \tag{4.9}$$

The point spread function associated with the recording microlens screen in the focal plane (which is the plane of the photographic plate) has an expression analogous to (4.6), but containing the parameters of the recording array:

$$K^{rec}_{sq,cl}(X,Y) = C\alpha^{rec}_{sq,cl} \exp\left(-\frac{X^2+Y^2}{\left(u^{rec}_{sq,cl}\right)^2}\right), \tag{4.10}$$

where $\alpha^{rec}_{sq,cl}$ and $u^{rec}_{sq,cl}$ are given by equations similar to (4.7) and (4.8), with $\phi := \phi_r$, $f := F$ - the pitch and the focal distance of a microlens in the recording screen, and $a := a_r$ - the distance between the reconstructed model P' and the recording microlens array ($a_r = D - |z_P|$).

The result of the entire optical process combines the spread effect of both the central double microlens screen and the recording array. Therefore, the spread function of the overall process is the convolution of $K^{ATS}_{sq,cl}$ and $K^{rec}_{sq,cl}$:

$$K^{total}_{sq,cl} = K^{ATS}_{sq,cl} \star K^{rec}_{sq,cl} \,. \tag{4.11}$$

Computations yield:

$$K_{sq,cl}^{total}(X,Y) = C\left(\alpha_{sq,cl}^{ATS}\right)^2 \alpha_{sq,cl}^{rec} \frac{\left(u_{sq,cl}^{ATS}\right)^2 u_{sq,cl}^{rec}}{\sqrt{2\left(u_{sq,cl}^{ATS}\right)^2 + \left(u_{sq,cl}^{rec}\right)^2}} \times \exp\left(-\frac{X^2+Y^2}{2\left(u_{sq,cl}^{ATS}\right)^2 + \left(u_{sq,cl}^{rec}\right)^2}\right). \tag{4.12}$$

Equation (4.12) expresses the spread function in local coordinates associated with a microlens of the recording screen in its focal plane. Using the global Cartesian system depicted in figure 4.1, the point spread behind a specific microlens k of the recording array, imaging a point P can be written as:

$$K_{sq,cl}^{total,k}(x,y) = C\left(\alpha_{sq,cl}^{ATS}\right)^2 \alpha_{sq,cl}^{rec} \frac{\left(u_{sq,cl}^{ATS}\right)^2 u_{sq,cl}^{rec}}{\sqrt{2\left(u_{sq,cl}^{ATS}\right)^2 + \left(u_{sq,cl}^{rec}\right)^2}} \times \exp\left(-\frac{(x-x'_k)^2+(y-y'_k)^2}{2\left(u_{sq,cl}^{ATS}\right)^2 + \left(u_{sq,cl}^{rec}\right)^2}\right), \tag{4.13}$$

where x'_k, and y'_k are the coordinates of the recorded intensity distribution P'_k behind microlens k, given in the previous section.

Parallax information about any imaged point P is spread on the photographic plate as a set of intensity distributions recorded in a number of microlenses, the number being determined by the axial position of the optical reconstruction P'. The point spread function of the whole process is the sum of these disparate distributions:

$$K_{sq,cl}^{global}(x,y) = \sum_k K_{sq,cl}^{total,k}(x,y)\,, \tag{4.14}$$

where the summation is done with respect to the indices k of the microlenses involved in recording point P.

Remote imaging and square based microlenses

In the context of remote imaging, the point spread function computed by Burckhardt et al. ([14]) for a square based microlens in its focal plane is given by:

$$K_{sq,rm}(X,Y) = C\alpha_{sq,rm} \exp\left(-\frac{X^2+Y^2}{u_{sq,rm}^2}\right) \tag{4.15}$$

where:

$$\alpha_{sq,rm} = \frac{b^2}{4.16b^2 + \pi^2\lambda^2 a^4}\,, \tag{4.16}$$

and the coefficient:

$$u^2_{sq,rm} = 0.104\frac{\phi^2 f^2 b^2}{a^4} + 0.245\frac{\lambda^2 f^2}{\phi^2}\,, \tag{4.17}$$

Following the same calculation steps, the point spread function of the whole optical process behind one microlens yields:

$$\begin{aligned} K^{total}_{sq,rm}(X,Y) = \quad & C\left(\alpha^{ATS}_{sq,rm}\right)^2 \alpha^{rec}_{sq,rm} \frac{\left(u^{ATS}_{sq,rm}\right)^2 u^{rec}_{sq,rm}}{\sqrt{2\left(u^{ATS}_{sq,rm}\right)^2 + \left(u^{rec}_{sq,rm}\right)^2}} \\ & \times \exp\left(-\frac{X^2+Y^2}{2\left(u^{ATS}_{sq,rm}\right)^2 + \left(u^{rec}_{sq,rm}\right)^2}\right), \end{aligned} \tag{4.18}$$

where the coefficients $\alpha^{ATS}_{sq,rm}$, $\alpha^{rec}_{sq,rm}$, $u^{ATS}_{sq,rm}$, and $u^{rec}_{sq,rm}$ are given by equations similar to (4.16) and (4.17) with parameters of the ATS and recording screen, respectively.

The global point spread function for a square based microlens camera configuration will be given, for both close and remote imaging, by the sum of local point spread functions behind all microlenses which image the considered quasi-pointwise light source as in equation (4.14).

4.3.2 Point spread function for circular based microlens integral system

Close imaging and circular based microlenses

Initially, the point spread function for a camera configuration based on circular based microlenses is derived, assuming close imaging conditions exist for both image formation and recording. The plane where the spread function is to be considered is the focal plane OXY of the microlens (Figure 3.11), where c denotes the image distance of the point source and f the focal distance of the microlens. The amplitude of the light wave in the focal plane is given by the Fresnel-Kirchhoff integral equation ([45], [14]):

$$\mathcal{A}(X,Y) = C \iint_{\mathcal{D}} \exp\left(\frac{ik}{2c}(x^2+y^2)\right) exp(-ikR)dx\,dy, \tag{4.19}$$

where $\mathcal{D} = \{(x,y) \mid x^2+y^2 \leq r^2\}$ is the lens aperture, $R = f^2+(X-x)^2+(Y-y)^2$ is the distance between a point in the plane of the lens and a point in the focal plane, and $k = 2\pi/\lambda$ is the wavenumber. The Taylor expansion and the thin lens formula are used to derive the following first order approximation, giving ([57]):

$$\begin{aligned}\mathcal{A}(X,Y) = \; & C\exp\left(-\frac{ik}{2f}(X^2+Y^2)\right)\iint_{\mathcal{D}}\exp\left(-\frac{ik}{2a}(x^2+y^2)\right) \\ & \times\exp\left(\frac{ik}{f}(xX+yY)\right)dx\,dy\,,\end{aligned} \tag{4.20}$$

where the half-depth b of the object P has been approximated to the distance a of P to the microlens, given the close imaging conditions.

Changing the variables to:

$$\hat{x} = -\frac{x}{\lambda f}, \quad \hat{y} = -\frac{y}{\lambda f}, \quad \hat{r} = \frac{r}{\lambda f}, \tag{4.21}$$

the equation becomes:

$$\begin{aligned}\mathcal{A}(X,Y) = \; & C\exp\left(-\frac{ik}{2f}(X^2+Y^2)\right)\iint_{\hat{\mathcal{D}}}\exp\left(-\frac{ik\lambda^2 f^2}{2a}(\hat{x}^2+\hat{y}^2)\right) \\ & \times\exp\left(-i2\pi(\hat{x}X+\hat{y}Y)\right)d\hat{x}\,d\hat{y}\,,\end{aligned} \tag{4.22}$$

where $\hat{\mathcal{D}}$ is the disk $\{(\hat{x},\hat{y}) \mid \hat{x}^2+\hat{y}^2 \leq \hat{r}^2\}$.

It can be noticed that the latter integral represents the Fourier transform of the product between $\psi_1 = \exp\left(-(ik\lambda^2 f^2/2a)(\hat{x}^2+\hat{y}^2)\right)$ and the characteristic function ψ_2 of the disk $\hat{\mathcal{D}}$, therefore the amplitude $\mathcal{A}$ is the convolution of the Fourier transforms of ψ_1 and ψ_2 ([57]):

$$\mathcal{A}(X,Y) = C\exp\left(-\frac{ik}{2f}(X^2+Y^2)\right)\mathcal{F}(\psi_1)\star\mathcal{F}(\psi_2)\,. \tag{4.23}$$

The two Fourier transforms are:

$$\mathcal{F}(\psi_1) = \frac{a}{\lambda f^2}\exp\left(-\frac{i\pi a}{\lambda f^2}(X^2+Y^2)\right), \tag{4.24}$$

and

$$\mathcal{F}(\psi_2) = C\frac{J_1(2\pi\rho\hat{r})}{2\pi\rho\hat{r}}\,, \tag{4.25}$$

respectively, where $\rho^2 = X^2 + Y^2$, and J_1 is the first order Bessel function. In order to perform the convolution, the approximation $2J_1(w)/w \simeq \exp(-w^2/6.5)$ is considered.

Further calculations yield ([57]):

$$\begin{aligned}\mathcal{A}(\rho) = \; & C\frac{a}{\sqrt{6.0736\hat{r}^2 + i\pi a/\lambda f^2}}\exp\left(-\frac{ik}{2f}\rho^2\right) \\ & \times\exp\left(-\beta\rho^2\left(\frac{\pi a}{\lambda f^2} + i\,6.0736\hat{r}^2\right)\right),\end{aligned} \tag{4.26}$$

where

$$\beta = \frac{6.0736\hat{r}^2\pi a\lambda f^2}{36.89\lambda^2 f^4\hat{r}^4 + \pi^2 a^2}\,. \tag{4.27}$$

Hence the point spread function is given by ([57]):

$$\begin{aligned}K_{circ,cl}(X,Y) &= \mathcal{A}(X,Y)\cdot\mathcal{A}^*(X,Y) \\ &= C\frac{a^2}{\sqrt{36.89\hat{r}^4 + \pi^2 a^2/\lambda^2 f^4}}\exp\left(-2\beta\frac{\pi a}{\lambda f^2}(X^2+Y^2)\right).\end{aligned} \tag{4.28}$$

Denoting the microlens diameter by $d = 2r$, the point spread function corresponding to a circular microlens yields ([57]):

$$K_{circ,cl}(X,Y) = C\alpha_{circ,cl}\exp\left(-\frac{X^2+Y^2}{u^2_{circ,cl}}\right), \tag{4.29}$$

where:

$$\alpha_{circ,cl} = \frac{a^2\lambda^2 f^2}{\sqrt{2.305d^4 + \pi^2 a^2\lambda^2}} \tag{4.30}$$

and

$$u^2_{circ,cl} = 0.076\frac{d^2 f^2}{a^2} + 0.328\frac{\lambda^2 f^2}{d^2}\,. \tag{4.31}$$

In Equation(4.31), the first term of the sum expresses the defocus influence, whilst the second expresses the diffraction. In addition, both $\alpha_{circ,cl}$ and $u_{circ,cl}$

contain depth information about the imaged point ([57]).

The calculation of the point spread function of the two tier integral imaging system shown in figure 4.1 has to take into account all three microlens arrays which light crosses prior to the integral reconstruction of the imaged object and subsequent recording ([57]). In the following analysis, it has been assumed that the spread influence of the macrolenses is negligible. The point spread function of the first microlens array in the focal plane is given by (4.29), where $a = a_{ATS} = (-z_P f_1 + z_P S + S^2)/(f_1 + z_P + S) - f$, d is the diameter, and f is the focal distance of a microlens in the ATS. It depends both upon the camera parameters and on the position of the point considered in space. Due to the symmetry of the optical process that occurs at the central double microlens array, the point spread function of the second microlens array is also given by (4.29). Therefore, the spread function of the spatial inversion process (associated with the central screen) is given by the convolution of (4.29) with itself:

$$K_{circ,cl}^{ATS} = K_{circ,cl}^{ATS} \star K_{circ,cl}^{ATS} \,. \tag{4.32}$$

Further computations yield ([57]):

$$K_{circ,cl}^{ATS}(X,Y) = C\left(\alpha_{circ,cl}^{ATS}\right)^2 u_{circ,cl}^{ATS} \exp\left(-\frac{X^2+Y^2}{2\left(u_{circ,cl}^{ATS}\right)^2}\right) . \tag{4.33}$$

The point spread function associated with the recording microlens screen in the focal plane (which is the plane of the photographic plate) is analogous to (4.29), but containing different parameters, and is given by ([57]):

$$K_{circ,cl}^{rec}(X,Y) = C\alpha_{circ,cl}^{rec} \exp\left(-\frac{X^2+Y^2}{\left(u_{circ,cl}^{rec}\right)^2}\right) , \tag{4.34}$$

where $\alpha_{circ,cl}^{rec}$ and $u_{circ,cl}^{rec}$ have expressions similar to (4.30) and (4.31), where $d := d_r$, $f := F$ are the diameter and the focal distance of the recording microlens screen respectively, and $a := a_r$ is the distance between the reconstructed model P' and the recording microlens array ($a_r = D - |z_P|$).

The spread function of the entire optical process combines the spread effect of both the central double microlens screen and the recording array. Therefore, it is

the convolution of $K^{ATS}_{circ,cl}$ and $K^{rec}_{circ,cl}$:

$$K^{total}_{circ,cl} = K^{ATS}_{circ,cl} \star K^{rec}_{circ,cl} \,. \tag{4.35}$$

Computations yield ([57]):

$$\begin{aligned} K^{total}_{circ,cl}(X,Y) = \quad & C \left(\alpha^{ATS}_{circ,cl}\right)^2 \alpha^{rec}_{circ,cl} \frac{\left(u^{ATS}_{circ,cl}\right)^2 u^{rec}_{circ,cl}}{\sqrt{2\left(u^{ATS}_{circ,cl}\right)^2 + \left(u^{rec}_{circ,cl}\right)^2}} \\ & \times \exp\left(-\frac{X^2+Y^2}{2\left(u^{ATS}_{circ,cl}\right)^2 + \left(u^{rec}_{circ,cl}\right)^2}\right). \end{aligned} \tag{4.36}$$

Equation (4.36) expresses the spread function in local coordinates associated with the microlens recording screen. Using the global Cartesian system depicted in figure 4.1, the point spread behind a specific microlens k of the recording array, imaging a point P can be written as ([57]):

$$\begin{aligned} K^{total,k}_{circ,cl}(x,y) = \quad & C \left(\alpha^{ATS}_{circ,cl}\right)^2 \alpha^{rec}_{psh,cl} \frac{\left(u^{ATS}_{circ,cl}\right)^2 u^{rec}_{circ,cl}}{\sqrt{2\left(u^{ATS}_{circ,cl}\right)^2 + \left(u^{rec}_{circ,cl}\right)^2}} \\ & \times \exp\left(-\frac{(x-x'_k)^2+(y-y'_k)^2}{2\left(u^{ATS}_{circ,cl}\right)^2 + \left(u^{rec}_{circ,cl}\right)^2}\right), \end{aligned} \tag{4.37}$$

where x'_k, and y'_k are the coordinates of the recorded intensity distribution P'_k behind microlens k, given in the previous section.

Parallax information about any imaged point P is distributed over the photographic plate as a set of intensity distributions recorded in a number of microlenses, the number being determined by the axial position of the optical reconstruction P'. The point spread function of the whole process is the sum of these disparate distributions ([57]):

$$K^{global}_{circ,cl}(x,y) = \sum_k K^{total,k}_{circ,cl}(x,y) \,, \tag{4.38}$$

where the summation is carried out with respect to the indices k of the microlenses

involved in recording point P.

Remote imaging and circular based microlenses

When *remote imaging* is assumed, a similar procedure for deriving the point spread function is used. The laws of geometrical optics show that the image distance is obtained as ([45], [13]):

$$c = f + \frac{f^2}{a^2} b ,$$

where a is the distance between object and lens. Using this equation, the Fresnel-Kirchhoff integral(4.19) becomes:

$$\mathcal{A}_{circ,rm}(X,Y) = C \iint_D \exp\left(-\frac{ik}{2a^2}(x^2+y^2)\right) exp\left(\frac{ik}{f}(xX+yY)\right) dx\, dy\,. \quad (4.39)$$

Computations similar to those performed in the close imaging case lead to the following expression for the point spread function for a circular microlens ([57]):

$$K_{circ,rm}(X,Y) = C\alpha_{circ,rm} \exp\left(-\frac{X^2+Y^2}{u^2_{circ,rm}}\right) , \quad (4.40)$$

where:

$$\alpha_{circ,rm} = \frac{\lambda^2 a^4 f^2}{b\sqrt{2.305 d^4 b^2 + \pi^2 \lambda^6 f^4 a^4}} , \quad (4.41)$$

and:

$$u^2_{circ,rm} = 0.076\frac{d^2 f^2 b^2}{a^4} + 0.328\frac{\lambda^2 f^2}{d^2} . \quad (4.42)$$

The point spread due to the whole optical process in the focal plane of a recording microlens k expressed in local coordinates attached to the microlens is ([57]):

$$\begin{aligned} K^{total,k}_{circ,rm}(X,Y) = \; & C\left(\alpha^{ATS}_{circ,rm}\right)^2 \alpha^{rec}_{circ,rm} \frac{\left(u^{ATS}_{circ,rm}\right)^2 u^{rec}_{circ,rm}}{\sqrt{2\left(u^{ATS}_{circ,rm}\right)^2 + \left(u^{rec}_{circ,rm}\right)^2}} \\ & \times \exp\left(-\frac{X^2+Y^2}{2\left(u^{ATS}_{circ,rm}\right)^2 + \left(u^{rec}_{circ,rm}\right)^2}\right) , \end{aligned} \quad (4.43)$$

where $\alpha^{ATS}_{circ,rm}$, $\alpha^{rec}_{circ,rm}$, and $u^{ATS}_{circ,rm}$, $u^{rec}_{circ,rm}$ are given by expressions analogous to (4.41) and (4.42) and the global point spread function for remote imaging conditions:

$$K^{global}_{circ,rm}(x,y) = \sum_k K^{total,k}_{circ,rm}(x,y)\,, \tag{4.44}$$

is analogous to (4.38).

4.3.3 Point spread function for hexagonal based microlens integral system

Analogous analyses can be made in order to derive the point spread function for camera configurations which involve hexagonal microlenses in both close and remote imaging contexts ([57]).

Close imaging and hexagonal based microlenses

Considering a hexagonal microlens of side l, computations similar to those performed for circular aperture microlenses in close imaging conditions give the following expression of the amplitude ([57]):

$$\mathcal{A}_{hex}(X,Y) = C\exp\left(-\frac{ik}{2f}(X^2+Y^2)\right)\mathcal{F}(\psi_1)\star\mathcal{F}(\psi_2)\,, \tag{4.45}$$

where $\psi_1 = \exp\left(-(ik\lambda^2 f^2/(2a))(\hat{x}^2+\hat{y}^2)\right)$, therefore:

$$\mathcal{F}(\psi_1) = \frac{a}{\lambda f^2}\exp\left(-\frac{i\pi a}{\lambda f^2}(X^2+Y^2)\right)\,. \tag{4.46}$$

The salient difference when comparing with equation (4.23) is that ψ_2 in this case represents the Fourier transform of the characteristic function of a hexagonal domain of side $\hat{l} = l/\lambda f$, f being the focal distance of a microlens ([57]):

$$\psi_2(\hat{x},\hat{y}) = \begin{cases} 1\,, & \text{if } (\hat{x},\hat{y}) \in Hex(\hat{l}) \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{4.47}$$

The Fourier transform of this hexagonal domain is real due to the symmetry of

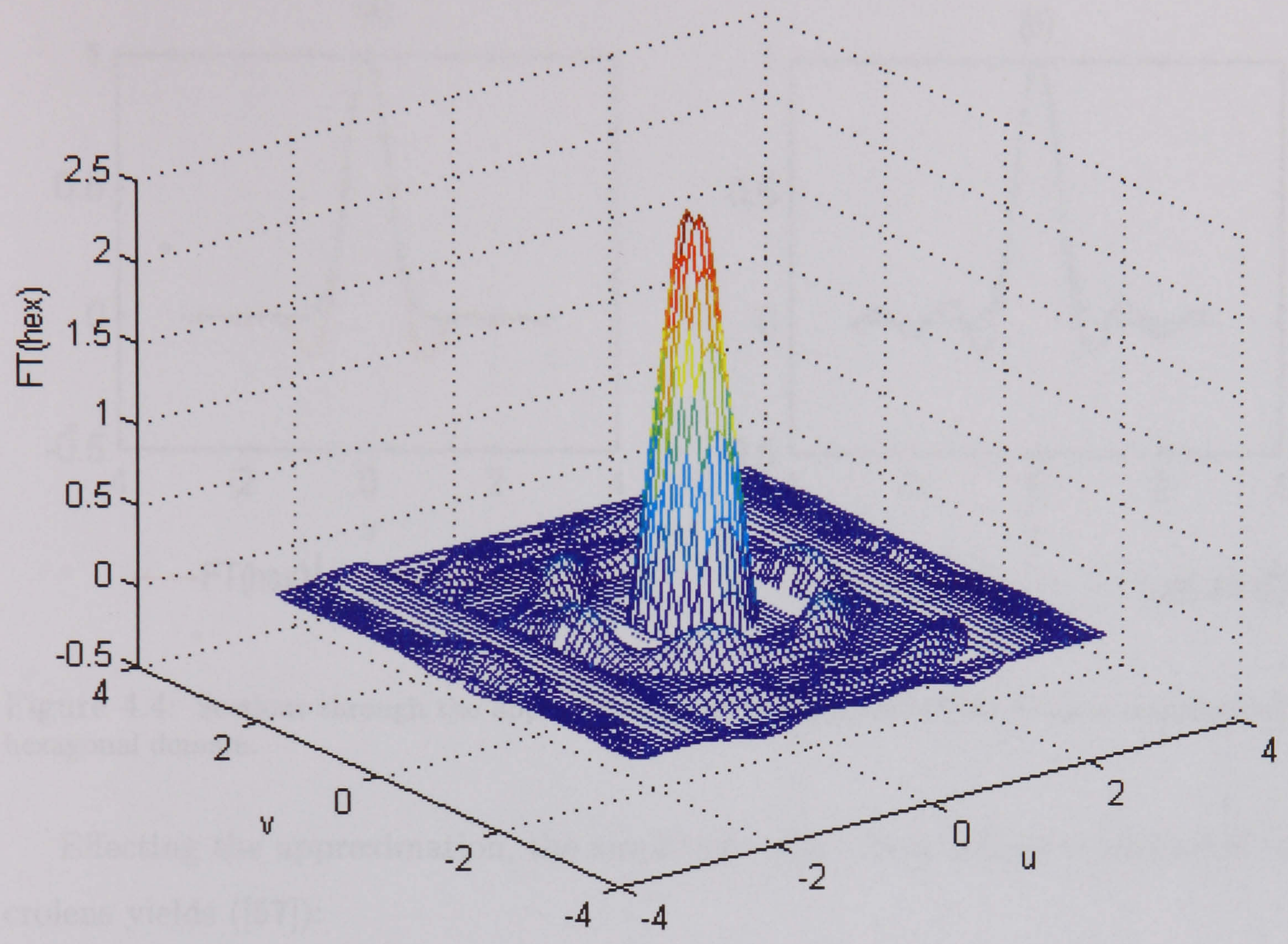


Figure 4.3: The Fourier transform of the characteristic function of a hexagonal domain.

ψ_2 and is given by ([57]):

$$\begin{aligned}\mathcal{F}(\psi_2) = \quad & \frac{1}{2\pi^2 Y(X + Y/\sqrt{3})}\left(\cos(\pi\hat{l}(Y - X\sqrt{3})) - \cos(2\pi\hat{l}Y)\right) \\ & + \frac{1}{2\pi^2 Y(X - Y/\sqrt{3})}\left(\cos(2\pi\hat{l}Y) - \cos(\pi\hat{l}(Y + X\sqrt{3}))\right).\end{aligned} \tag{4.48}$$

More computational details can be found in Appendix B.

In order to perform the convolution, an approximation to this particular Fourier transform is necessary. The graphical representation of (4.48), shown in Figure 4.3, which has the shape of an asymmetric bell, suggests a Gaussian approximation which results in ([57]):

$$\mathcal{F}(\psi_2) \simeq \frac{3\sqrt{3}}{2}\hat{l}^2 \exp\left(-4.8\hat{l}^2X^2 - 5.3\hat{l}^2Y^2\right). \tag{4.49}$$

Figure 4.4 shows how the chosen Gaussian approximates the Fourier transform of a hexagonal domain.

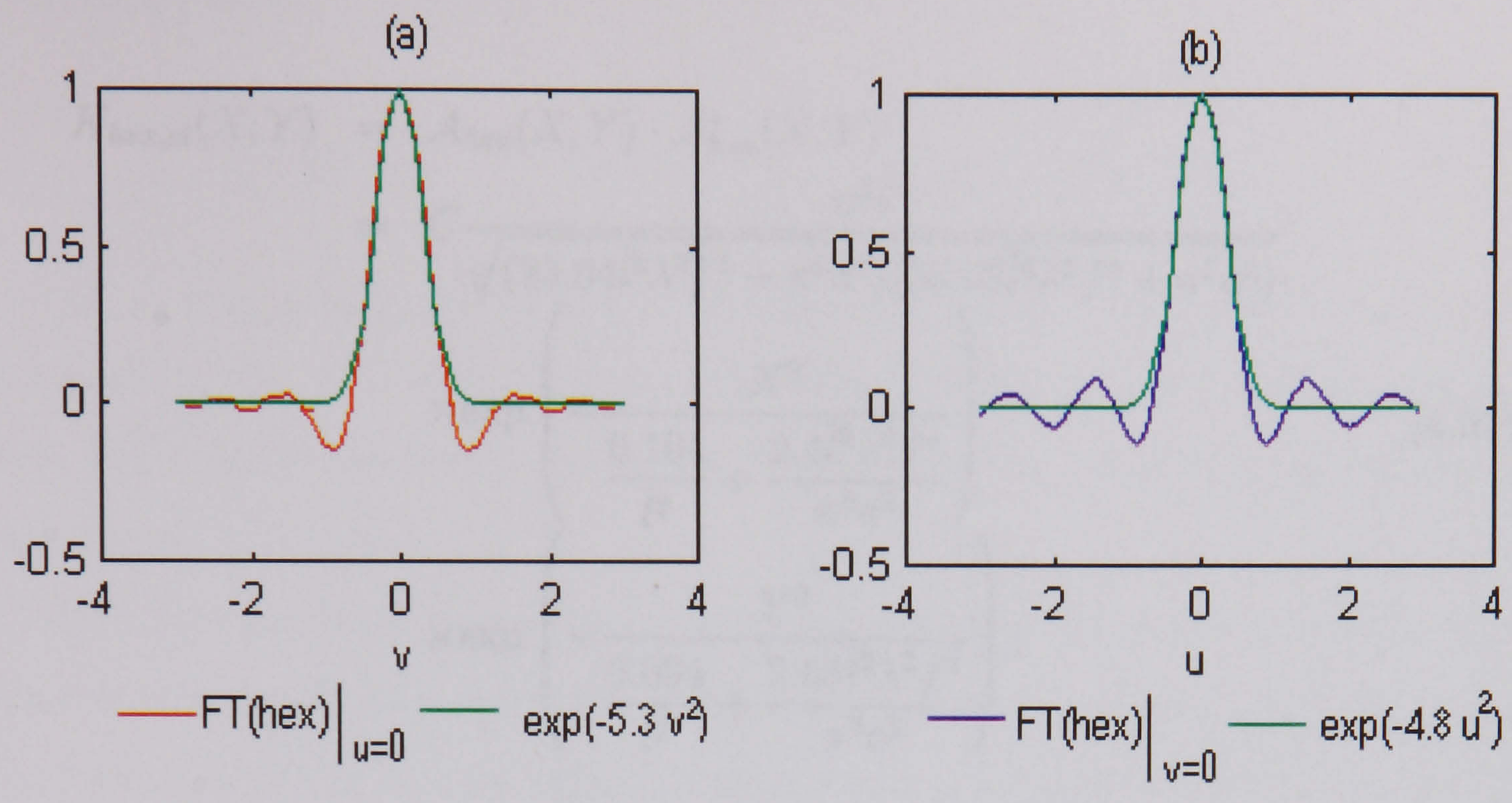


Figure 4.4: Sections through the approximations with Gaussians of the Fourier transform of a hexagonal domain.

Effecting the approximation, the amplitude of the light behind a hexagonal microlens yields ([57]):

$$\begin{aligned}\mathcal{A}_{hex}(X,Y) \;=\; C\;\; & \frac{a\hat{l}^2}{\lambda f^2}\;\; \frac{\exp\left(-\frac{ik}{2f}(X^2+Y^2)\right)}{\sqrt{4.8\hat{l}^2+i\pi a/(\lambda f^2)}\sqrt{5.3\hat{l}^2+i\pi a/(\lambda f^2)}} \\ & \times\exp\left(-X^2\frac{4.8\pi^2a^2\hat{l}^2+i4.8^2\lambda f^2\hat{l}^4}{\pi^2a^2+4.8^2\lambda^2f^4\hat{l}^4}\right) \\ & \times\exp\left(-Y^2\frac{5.3\pi^2a^2\hat{l}^2+i5.3^2\lambda f^2\hat{l}^4}{\pi^2a^2+5.3^2\lambda^2f^4\hat{l}^4}\right),\end{aligned} \tag{4.50}$$

therefore the point spread function of a hexagonal (micro)lens under close imaging conditions, which is the product between the amplitude and its complex conjugate, is given by ([57]):

$$
\begin{aligned}
K_{hex,cl}(X,Y) &= \mathcal{A}_{hex}(X,Y)\cdot\mathcal{A}^{*}_{hex}(X,Y) \\
&= C\frac{a^2\hat{l}^2}{\sqrt{(23.04\hat{l}^4\lambda^2f^4+\pi^2a^2)(28.09\hat{l}^4\lambda^2f^4+\pi^2a^2)}} \\
&\times\exp\left(-\frac{X^2}{\frac{0.104}{\hat{l}^2}+\frac{2.4\hat{l}^2\lambda^2f^4}{\pi^2a^2}}\right) \\
&\times\exp\left(-\frac{Y^2}{\frac{0.094}{\hat{l}^2}+\frac{2.65\hat{l}^2\lambda^2f^4}{\pi^2a^2}}\right),
\end{aligned}
\tag{4.51}
$$

When replacing $\hat{l}$ by $l/(\lambda f)$, the following equation is obtained ([57]):

$$
K_{hex,cl}(X,Y) = C\alpha^{ATS}_{hex,cl}\exp\left(-\frac{X^2}{(u^{ATS}_{hex,cl})^2}-\frac{Y^2}{(v^{ATS}_{hex,cl})^2}\right), \tag{4.52}
$$

with:

$$
\alpha^{ATS}_{hex,cl} = \frac{a^2l^2}{f^2\sqrt{(23.04l^4+\pi^2a^2\lambda^2)(28.09l^4+\pi^2a^2\lambda^2)}} \tag{4.53}
$$

and:

$$
(u^{ATS}_{hex,cl})^2 = 0.243\frac{l^2f^2}{a^2}+0.104\frac{\lambda^2f^2}{l^2}, \tag{4.54}
$$

$$
(v^{ATS}_{hex,cl})^2 = 0.270\frac{l^2f^2}{a^2}+0.094\frac{\lambda^2f^2}{l^2}, \tag{4.55}
$$

where $a = a_{ATS} = (-z_Pf_1+z_PS+S^2)/(f_1+z_P+S)-f$ is the distance between the point and the central screen.

The spread associated with the central double microlens screen ATS in this case is the convolution of (4.52) with itself, given the symmetry of the arrangement. Hence ([57]):

$$
\begin{aligned}
K^{ATS}_{hex,cl} &= K_{hex,cl}\star K_{hex,cl} \\
&= C(\alpha^{ATS}_{hex,cl})^2u^{ATS}_{hex,cl}v^{ATS}_{hex,cl}\exp\left(-\tfrac{X^2}{2(u^{ATS}_{hex,cl})^2}\right)\exp\left(-\tfrac{Y^2}{2(v^{ATS}_{hex,cl})^2}\right)
\end{aligned}
\tag{4.56}
$$

Similarly to (4.52), the point spread associated with the recording microlenses is

obtained as ([57]):

$$K^{rec}_{hex,cl}(X,Y) = C\alpha^{rec}_{hex,cl} \exp\left(-\frac{X^2}{(u^{rec}_{hex,cl})^2} - \frac{Y^2}{(v^{rec}_{hex,cl})^2}\right), \tag{4.57}$$

with $\alpha^{rec}_{hex,cl}$, $u^{rec}_{hex,cl}$, $v^{rec}_{hex,cl}$ given by expressions analogous to (4.53)-(4.54), where the parameters of the ATS have been replaced with the parameters of the recording screen, and a is the distance between the reconstructed model P' and the recording array: $a = a_r = D - |z_P|$.

Using these expressions and the fact that the spread of the entire optical process is the convolution of the spreads due to each of the microlens arrays, the total point spread function of the camera using hexagonal microlenses yields ([57]):

$$K^{total,k}_{hex,cl}(X,Y) = C\left(\alpha^{ATS}_{hex,cl}\right)^2 \alpha^{rec}_{hex,cl}\Gamma_{hex,cl} \exp\left(-\frac{X^2}{U^2_{hex,cl}} - \frac{Y^2}{V^2_{hex,cl}}\right) \tag{4.58}$$

where

$$\Gamma_{hex,cl} = \frac{\left(u^{ATS}_{hex,cl}\right)^2\left(v^{ATS}_{hex,cl}\right)^2 u^{rec}_{hex,cl}v^{rec}_{hex,cl}}{\sqrt{\left(2\left(u^{ATS}_{hex,cl}\right)^2+\left(u^{rec}_{hex,cl}\right)^2\right)\left(2\left(v^{ATS}_{hex,cl}\right)^2+\left(v^{rec}_{hex,cl}\right)^2\right)}}, \tag{4.59}$$

$$U^2_{hex,cl} = 2\left(u^{ATS}_{hex,cl}\right)^2+\left(u^{rec}_{hex,cl}\right)^2, \tag{4.60}$$

$$V^2_{hex,cl} = 2\left(v^{ATS}_{hex,cl}\right)^2+\left(v^{rec}_{hex,cl}\right)^2, \tag{4.61}$$

and the α-, u-, and v-factors are given by equations (4.53) and (4.54) whose parameters have been adapted for the corresponding microlens screen, as indicated by the indices rec and ATS.

The point spread function of the whole process is the sum of the disparate intensity distributions recorded in a number of microlenses, depending on the axial position of the optical model:

$$K^{global}_{hex,cl} = \sum_k K^{total,k}_{hex,cl}(x,y), \tag{4.62}$$

where the summation is done with respect to the indices k of the microlenses involved in recording point P ([57]).

Remote imaging and hexagonal based microlenses

The derivation of the point spread function for the two tier integral camera system employing hexagonal based microlenses follows similar steps when remote imaging conditions are assumed ([57]).

The integration of the Fresnel-Kirchhoff equation over the hexagonal aperture of a microlens yields the following equation for the point spread behind one microlens ([57]):

$$K_{hex,rm} = C\alpha_{hex,rm} \exp\left(-\frac{X^2}{u_{hex,rm}^2}\right) \exp\left(-\frac{Y^2}{v_{hex,rm}^2}\right), \tag{4.63}$$

where:

$$\alpha_{hex,rm} = \frac{a^4 l^4 \lambda^2}{\sqrt{23.04l^4 + \pi^2\lambda^6 a^4 f^4}\sqrt{28.09l^4 + \pi^2\lambda^6 a^4 f^4}} \tag{4.64}$$

$$u_{hex,rm} = 0.243\frac{l^2 f^2 b^2}{a^4} + 0.104\frac{\lambda^2 f^2}{l^2} \tag{4.65}$$

$$v_{hex,rm} = 0.270\frac{l^2 f^2 b^2}{a^4} + 0.094\frac{\lambda^2 f^2}{l^2}. \tag{4.66}$$

The optical process has been followed step by step, and, the same as in the case of circular based microlenses, the spread corresponding to each microlens array has been computed. Then, the convolution of all these intermediary successive spreads has yielded the total point spread function behind a certain microlens of the recording array in remote imaging conditions ([57]):

$$K_{hex,rm}^{total,k}(X,Y) = C\left(\alpha_{hex,rm}^{ATS}\right)^2 \alpha_{hex,rm}^{rec} \Gamma_{hex,rm} \exp\left(-\frac{X^2}{U_{hex,rm}^2} - \frac{Y^2}{V_{hex,rm}^2}\right) \tag{4.67}$$

where $\Gamma_{hex,rm}$, $U_{hex,rm}$, and $V_{hex,rm}$ are:

$$\Gamma_{hex,rm} = \frac{\left(u_{hex,rm}^{ATS}\right)^2\left(v_{hex,rm}^{ATS}\right)^2 u_{hex,rm}^{rec} v_{hex,rm}^{rec}}{\sqrt{\left(2\left(u_{hex,rm}^{ATS}\right)^2 + \left(u_{hex,rm}^{rec}\right)^2\right)\left(2\left(v_{hex,rm}^{ATS}\right)^2 + \left(v_{hex,rm}^{rec}\right)^2\right)}}, \tag{4.68}$$

$$U_{hex,rm}^2 = 2\left(u_{hex,rm}^{ATS}\right)^2 + \left(u_{hex,rm}^{rec}\right)^2, \tag{4.69}$$

$$V_{hex,rm}^2 = 2\left(v_{hex,rm}^{ATS}\right)^2 + \left(v_{hex,rm}^{rec}\right)^2, \tag{4.70}$$

and $\alpha_{hex,rm}^{ATS}$, $\alpha_{hex,rm}^{rec}$, $u_{hex,rm}^{ATS}$, $u_{hex,rm}^{rec}$, $v_{hex,rm}^{ATS}$, $v_{hex,rm}^{rec}$ are given by (4.64) with appropriate parameters according to the microlens layer they refer to.

The global point spread function for a hexagonal microlens camera configuration is given by the sum of local point spread functions behind all microlenses which image the quasi-pointwise light source ([57]). The expression is simply:

$$K^{global}_{hex,rm}(x,y) = \sum_k K^{total,k}_{hex,rm}(x,y)\,, \tag{4.71}$$

4.4 Depth information encoded in the omnidirectional integral image

4.4.1 Number of recorded intensity distributions corresponding to a physical point

The number of recorded intensity distributions of the optical reconstruction P' of a physical point P (figure 4.2(b)) is important particularly with respect to the way information is distributed. The imaged point $P(x,y,z)$ is recorded behind a number of microlenses on the photographic plate such that the spatial and angular information thus distributed creates the continuous parallax in all directions visible at the 3D-replay ([57]).

This number depends not only on the spatial position of the point, but also on the types of microlens packaging in the recording array. There are two types of packaging:

- rectangular net packaging
- honeycomb packaging

as shown in Figures 4.5 (a)-(d).

The rectangular net packaging is used for square and circular based microlenses (Figures 4.5 (a)-(b)), and the honeycomb packaging can be used for circular and hexagonal based microlens arrays (Figures 4.5 (c)-(d)).

In the following, it is assumed that the aperture of the camera can be approximated to a square of side ap. The microlenses which 'see' and record intensity distributions pertaining to a physical point $P(x_P, y_P, z_P)$ have the x coordinate of

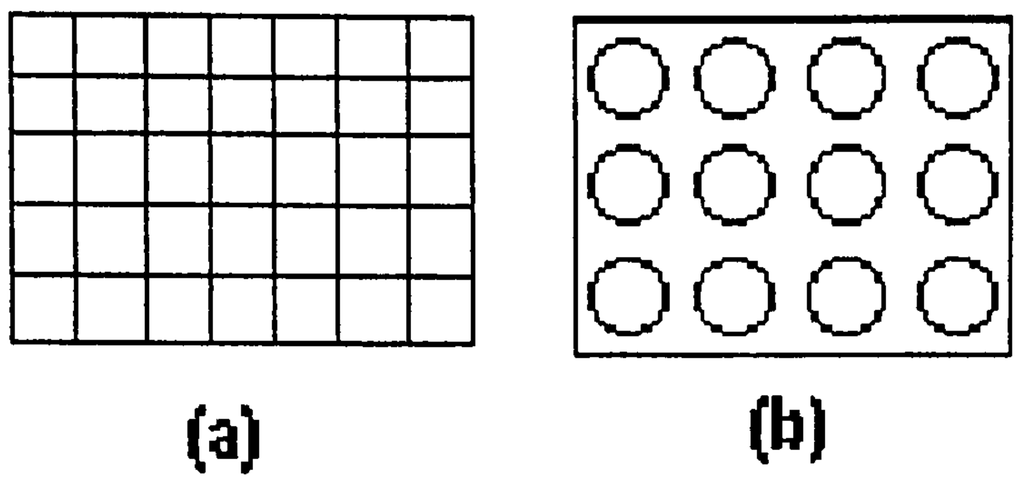


Figure 4.5: Packaging types: (a) rectangular net packaging for square based microlens arrays, (b) rectangular net packaging for circular based microlens arrays, (c) honeycomb packaging for circular based based microlens arrays, (d) honeycomb packaging for hexagonal based microlens arrays.

the centre in the interval:

$$D_x = \left[\max\left(\frac{ap}{2} + (x_P - \frac{ap}{2})\frac{D-S}{|z_P|-S}, -\frac{ap}{2}\right), \min\left(-\frac{ap}{2} + (x_P + \frac{ap}{2})\frac{D-S}{|z_P|-S}, \frac{ap}{2}\right)\right], \tag{4.72}$$

and the y coordinate in the interval:

$$D_y = \left[\max\left(ap + (y_P - ap)\frac{D-S}{|z_P|-S}, 0\right), \min\left(y_P\frac{D-S}{|z_P|-S}, ap\right)\right], \tag{4.73}$$

The maximum recorded image number for a given depth is attained for points having a central position in the imaged scene. This maximum has been derived for each of the packaging types ([57]).

For square or circular based microlenses packed in a rectangular net this maximum is ([57]):

$$N_{max}^{rn} = \left(\left[ap\frac{D-|z_P|}{\phi_r(|z_P|-S)}\right] + 1\right)^2, \tag{4.74}$$

whilst for circular or hexagonal based honeycomb packed microlens configurations (Figure 4.5(b),(c)), it is given by ([57]):

$$N_{max}^{hc} = \left(\left[ap\frac{D-|z_P|}{\frac{\sqrt{3}}{2}\phi_r(|z_P|-S)}\right] + 1\right)\left(\left[ap\frac{D-|z_P|}{\phi_r(|z_P|-S)}\right] + 1\right), \tag{4.75}$$

where ap is the side of the quasi-square camera aperture.

These computations show that this number depends on both the lateral position and depth of the point, as well as on the configuration of the camera (microlens pitch, microlens type and packaging) ([57]).

Comparing the results given in equations (4.74) and(4.75), it can be inferred that, given the same camera aperture, microlens pitch, and point position, the recording process using honeycomb packed microlenses retains more angular information than that corresponding to rectangularly packed microlenses. This is due to the fact that the honeycomb packaging reduces the effective microlens pitch ϕ_r to $\sqrt{3}\phi_r/2$ ([57]).

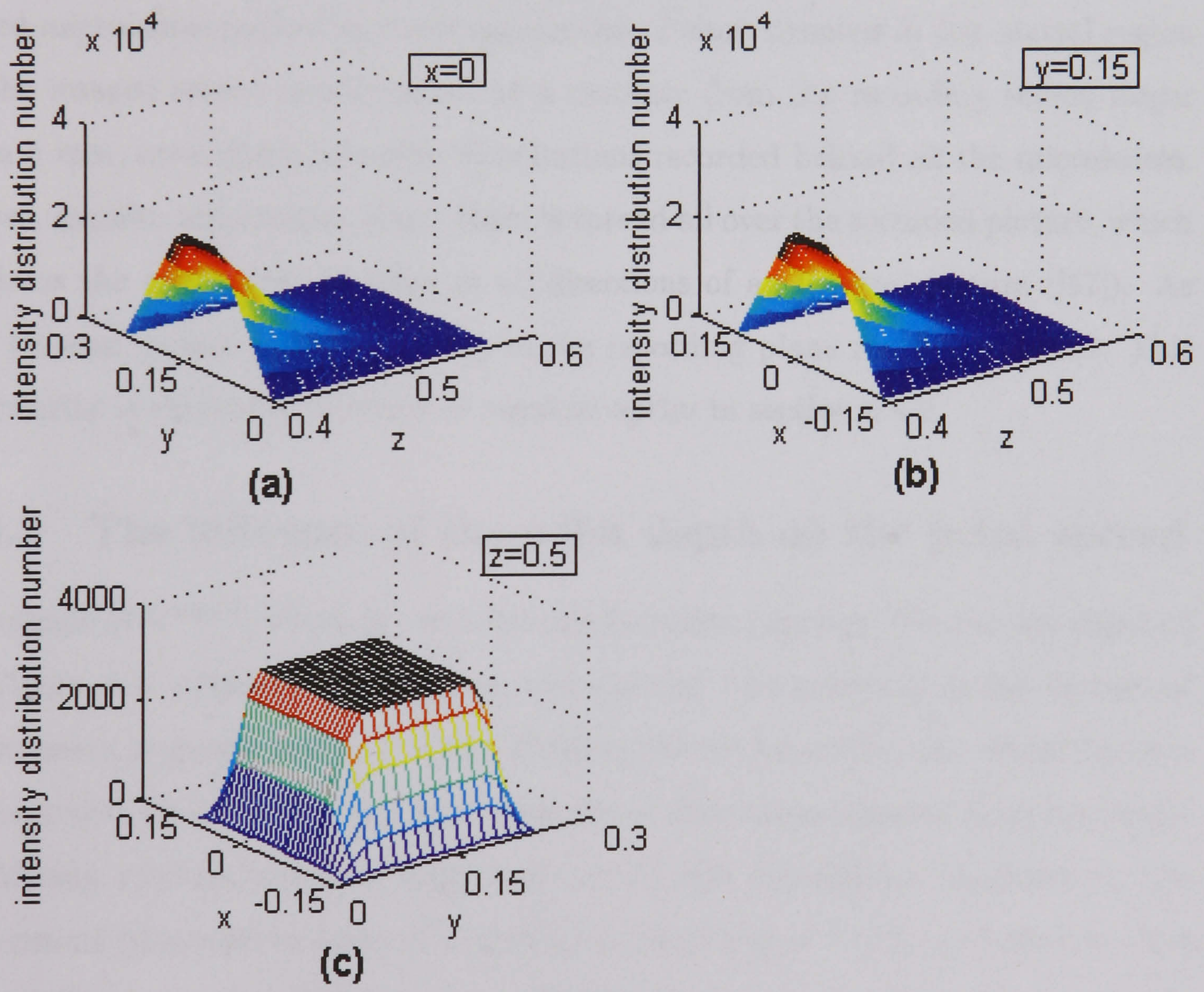


Figure 4.6: The number of recorded microintensity distributions for a hemi-spherical microlens system: (a) $x = const$, (b) $y = const$, (c) $z = const$.

It can be seen that the number of recorded intensity distribution components depends essentially on the camera aperture ap with marginal points having less encoded spatial distribution components due to the finite dimensions of the recording screen. A Matlab application has been used to plot the step functions that describe

the dependence between the number of recorded intensity distribution components of a point and the depth for different lateral positions of the point in Figures 4.6 (a), (b), (c). Figures 4.6(a)-(b) depict the number of recorded components for points in a plane parallel to Oyz and to Oxz respectively ([57]), with image plane depth ranging between 0.4 m and 0.58 m when the recording screen is at a distance D=0.58 m from the central plane of a camera having an aperture ap=0.3 m and honeycomb packed circular or hexagonal microlenses of pitch ϕ_r=1.65 mm in the recording array. The number of recorded intensity distributions for points having the same depth $|z| = z_0 = 0.50m$ and different lateral positions is represented in Figure 4.6(c). Similarly shaped graphs are obtained in the case of square or circular based microlenses packed in a rectangular net. Points situated in the central region of the imaged scene, reconstructed at a distance from the recording screen larger than 5 mm, have micro-intensity distributions recorded behind all the microlenses. Hence, angular information about them is spread all over the recorded picture, which explains the continuous parallax in all directions of an integral picture ([57]). As can be seen, points in the proximity of the recording plane are not recorded. This singularity is explained in terms of physical optics in section 4.4.2.

4.4.2 The influence of the point depth on the point spread

The shape of $K^{total,k}$, which is a two variable Gaussian function, like the one depicted in Figure 4.7, symmetric for square and circular and asymmetric for hexagonal microlenses, suggests that the energy forming the recorded intensity distribution is concentrated on a circular or elliptic spot whose dimensions depend on point depth.

Matlab applications have been used to plot the dependence on depth for the dimensions (diameter or axes) of a circular or elliptic spot which concentrates more than 98% of the recorded intensity. Both close and remote imaging conditions have been studied. Tables 4.1-4.4 and Figures 4.8-4.10 present spread dependence on depth in a camera configuration corresponding to a recording screen position D =0.585 m, at close imaging (Tables 4.1-4.2, Figure 4.8) and remote imaging (Tables 4.3-4.4, Figures 4.9-4.10). A wavelength λ =500 nm has been assumed. To make the comparison consistent, the same microlens pitches have been considered for all three microlens types, namely ϕ_{ATS}=0.125 mm for the microlenses in the central

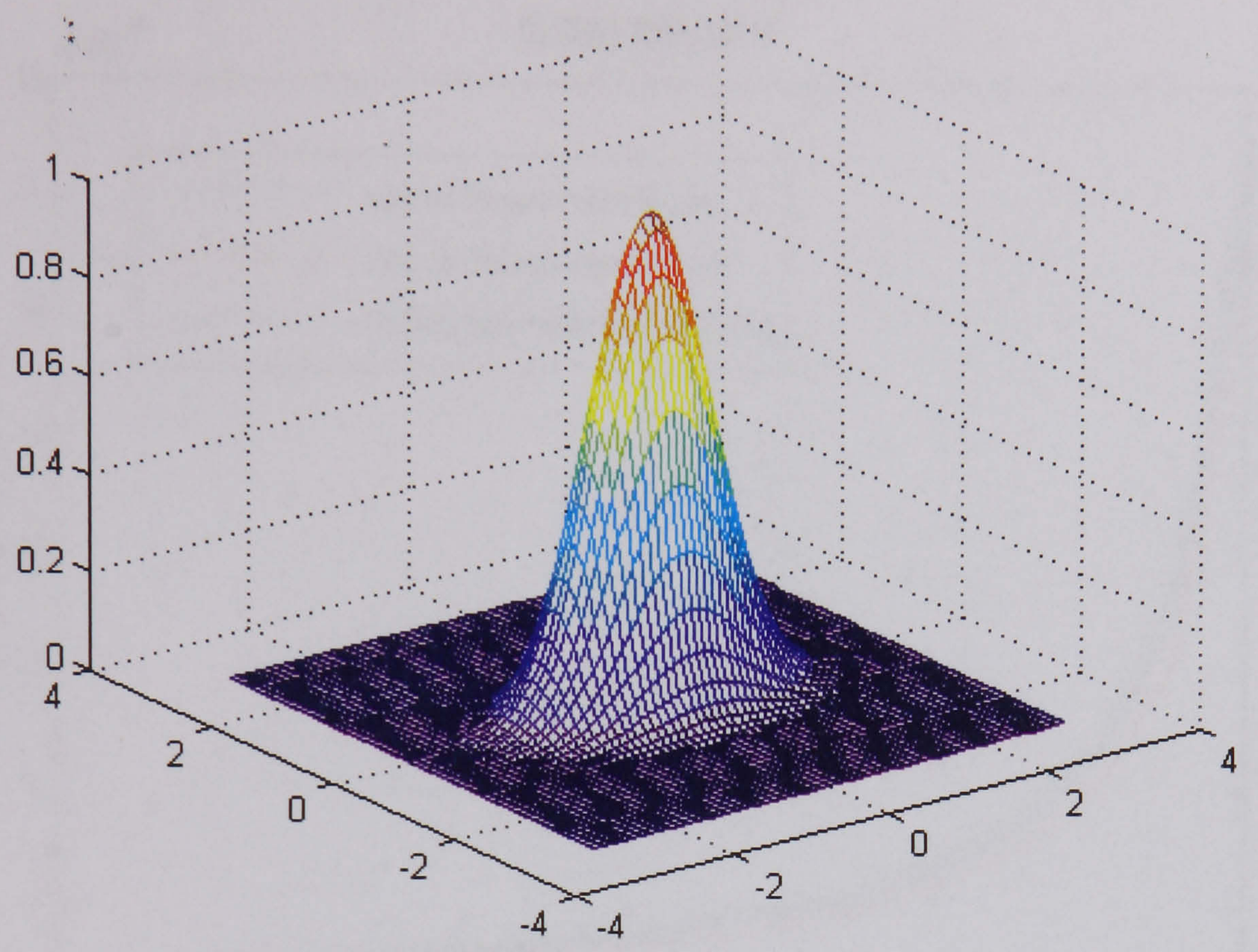


Figure 4.7: Graphical representation of a two dimensional asymmetric Gaussian.

double microlens screen, and ϕ_{rec}=1.206 mm for the microlenses in the recording screen. The tables show that the spread is larger when the recorded points are nearer the recording screen. For points reconstructed in the immediate vicinity of the recording plane, at a distance smaller than the focal distance of a recording microlens, the spread tends to infinity, therefore these points do not contribute stricto sensu to the recorded spatial information ([57]). This result confirms that the number of contributing intensity distributions for these points is zero, as proved on a geometrical basis in the last section. A continuous decrease of the spread is noticed when the distance from the optical reconstruction to the recording screen grows. Hence, the spread dimensions encode depth information about the imaged scene ([57]).

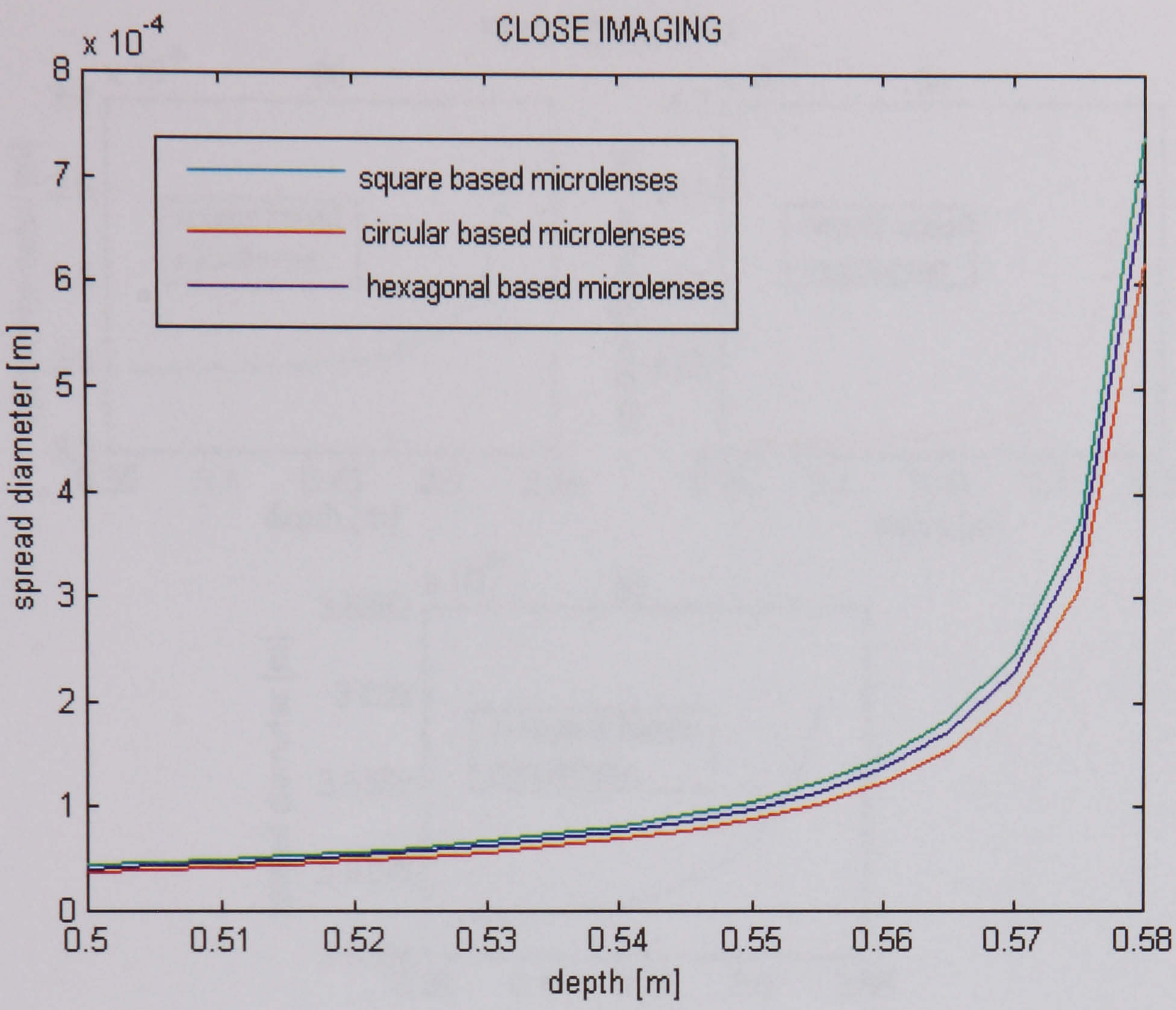


Figure 4.8: Comparison between the point spread as a function of depth for square, circular, and hexagonal microlens two tier integral systems at close imaging.

If close imaging conditions are assumed, the variations of the spread with respect to point depth corresponding to each microlens type are shown for comparison in Figure 4.8. The numerical details of the graphs are shown in Table 4.1, whilst Table 4.2 presents the ratios between spreads corresponding to different types of microlens systems as a function of depth in close imaging conditions. Making a comparative analysis between the three types of microlenses from this point of view, it can be observed that, in the case of close imaging, the lowest spread is obtained for circular based microlenses ([57]). It is approximately 16-17% lower than the spread obtained with square based microlenses, and 10% lower than the spread corresponding to hexagonal based microlenses. At their turn, hexagonal microlenses produce a spread 7% lower than that obtained with square based microlenses.

A similar analysis has been performed for remote imaging conditions. The graphs of spread versus depth for each microlens type are shown separately in Figures 4.9 and all together in Figure 4.10. Table 4.3 presents the actual spreads read in the graphs, whose mutual ratios are listed in Table 4.4. The lowest spread at remote

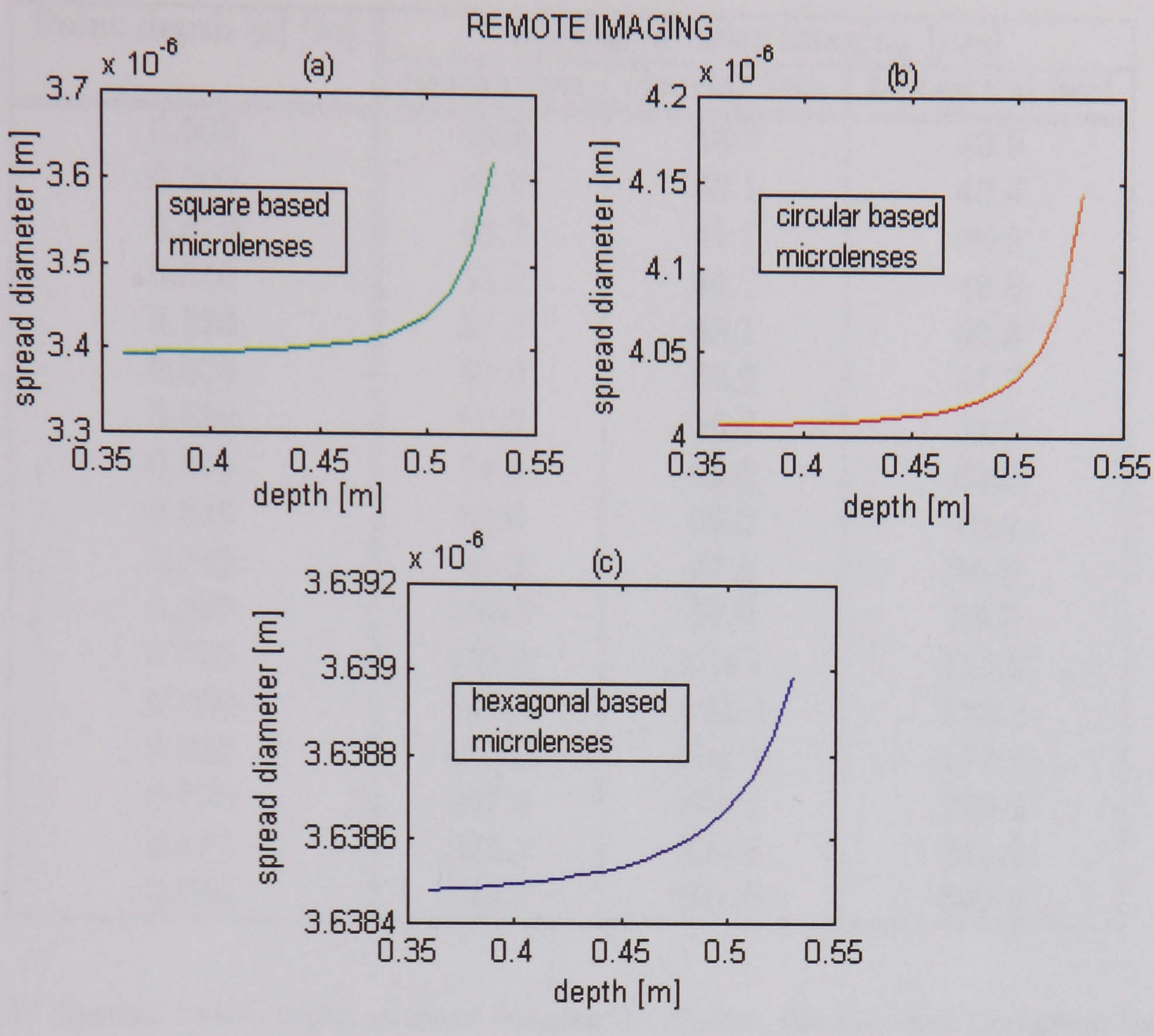


Figure 4.9: The point spread as a function of depth for: (a) square, (b) circular, and (c) hexagonal microlens two tier integral systems at remote imaging.

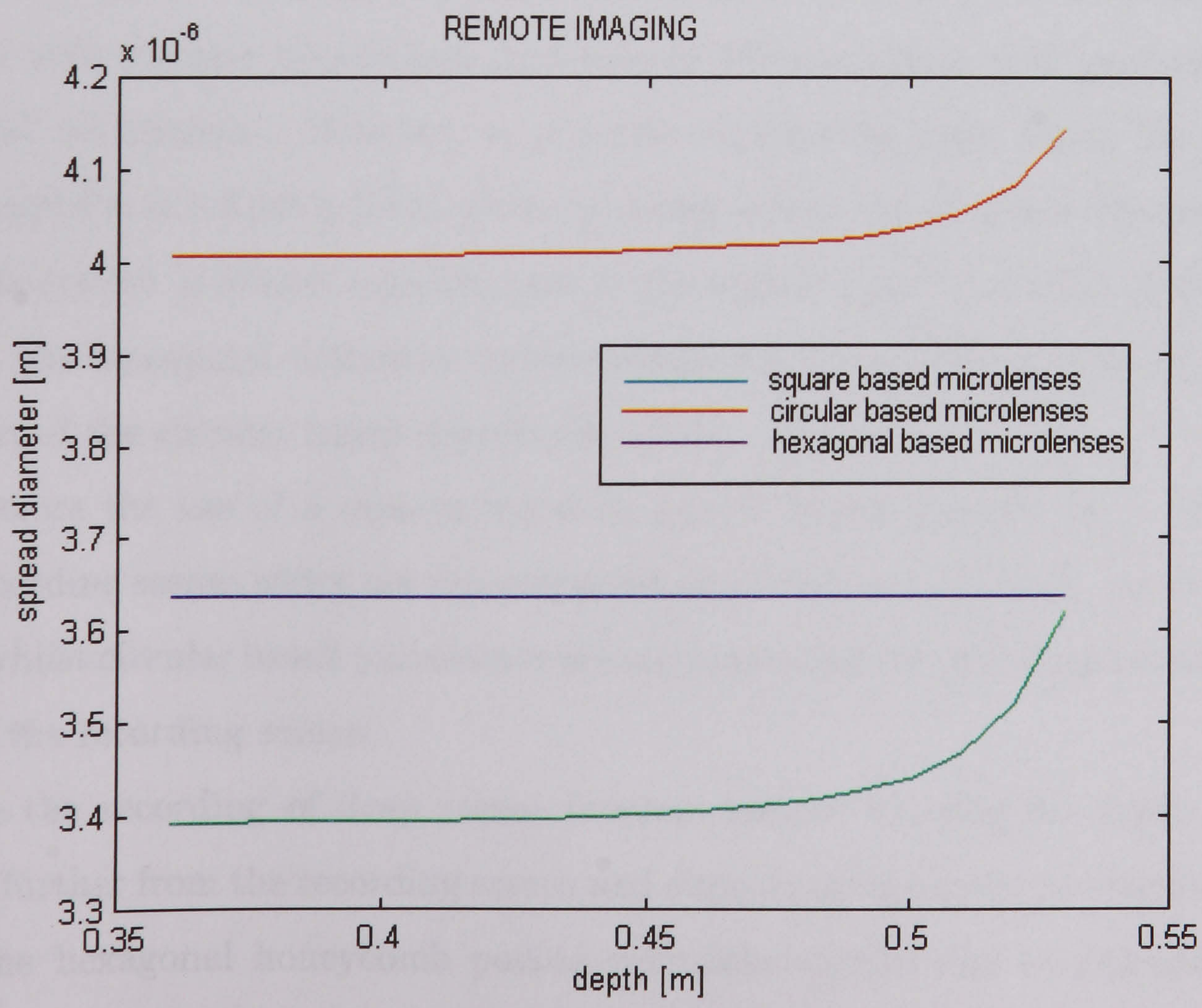


Figure 4.10: Comparison between the point spread as a function of depth for square, circular, and hexagonal microlens two tier integral systems at remote imaging.

Point depth $\|z\|$ [m]	Spread at close imaging $[\mu m]$		
	square lens	circular lens	hexagonal lens
0.500	43.9	36.9	40.9
0.505	46.6	39.1	43.4
0.510	49.7	41.7	46.3
0.515	53.2	44.7	49.6
0.520	57.3	48.1	53.3
0.525	62.0	52.0	57.7
0.530	67.6	56.7	63.0
0.535	74.4	62.3	69.2
0.540	82.6	69.2	76.9
0.545	92.9	77.8	86.4
0.550	106.1	88.9	98.7
0.555	123.8	103.7	115.2
0.560	148.5	124.4	138.2
0.565	185.6	155.4	172.7
0.570	247.4	207.2	230.2
0.575	371.1	310.8	345.2
0.580	742.1	621.5	690.4

Table 4.1: Spread versus depth at close imaging for square, circular and hexagonal based microlenses.

imaging is obtained with square based microlenses. It is 13-15% lower than that obtained with circular microlenses, and around 7% lower than that produced with hexagonal microlenses. However, it is worth mentioning that, when the imaged optical model is at 5-6 cm in front of the recording screen, the spread of the hexagonal microlens system is almost equal to that of the square based microlens system. At its turn, the hexagonal microlens system exhibits a spread which is 10-13% lower than that of the circular based microlenses ([57]).

Therefore the use of a camera based on square based microlenses is advisable when recording scenes which are reconstructed at a distance away from the recording screen, whilst circular based microlenses are recommended for recording scenes which straddle the recording screen.

Since the recording of deep scenes involves remote imaging for depth planes situated further from the recording screen and close imaging for the proximate depth range, the hexagonal honeycomb packed microlens system can be considered as presenting certain advantages in this situation. The behaviour of the hexagonal

Point depth $\|z\|$ [m]	Spread ratios at close imaging		
	$spread_{circ}/spread_{hex}$	$spread_{circ}/spread_{sq}$	$spread_{hex}/spread_{sq}$
0.500	90.22%	84.05%	93.17%
0.505	90.09%	83.91%	93.13%
0.510	90.06%	83.90%	93.16%
0.515	90.12%	84.02%	93.23%
0.520	90.24%	83.94%	93.02%
0.525	90.12%	83.87%	93.06%
0.530	90.00%	83.88%	93.20%
0.535	90.03%	83.74%	93.01%
0.540	89.99%	83.78%	93.10%
0.545	90.05%	83.75%	93.00%
0.550	90.07%	83.79%	93.03%
0.555	90.02%	83.76%	93.05%
0.560	90.01%	83.77%	93.06%
0.565	89.98%	83.73%	93.05%
0.570	90.01%	83.75%	93.05%
0.575	90.03%	83.75%	93.02%
0.580	90.02%	83.75%	93.03%

Table 4.2: Spread versus depth at close imaging for square, circular and hexagonal based microlenses. Ratios between spreads corresponding to the types of microlenses considered.

microlens system is average at both close and remote imaging. In close imaging conditions, the spread of the hexagonal system is lower than that of the square based microlenses, and in spite of having a higher spread than the circular based microlenses, the hexagonal system has the advantage of using more efficiently the surface of the microlens arrays. On a hexagonal microlens array, there are no inactive areas, while on a circular based microlens array significant interleavings are left between microlenses ([57]).

The hexagonal microlens system presents significant advantages at remote imaging, too. As already mentioned, it performs better than the circular based microlens one in terms of spread. Moreover, though the spread of the square based microlenses is lower in the remote imaging case, the hexagonal system has the additional advantage of retaining more angular information at a given pitch due to the honeycomb packaging, as opposed to the rectangular net packaging inherent to the square based microlens arrays ([57]), as discussed in Section 4.4.1.

Point depth $\|z\|$ [m]	Spread at remote imaging $[\mu m]$		
	square lens	circular lens	hexagonal lens
0.36	3.3929	4.0066	3.63847
0.37	3.3932	4.0068	3.63848
0.38	3.3937	4.0071	3.63848
0.39	3.3942	4.0074	3.63848
0.40	3.3949	4.0078	3.63849
0.41	3.3957	4.0083	3.63849
0.42	3.3968	4.0089	3.63850
0.43	3.3981	4.0097	3.63851
0.44	3.3999	4.0108	3.63852
0.45	3.4022	4.0121	3.63853
0.46	3.4053	4.0140	3.63855
0.47	3.4096	4.0166	3.63857
0.48	3.4159	4.0203	3.63859
0.49	3.4255	4.0260	3.63862
0.50	3.4409	4.0352	3.63867
0.51	3.4675	4.0512	3.63873
0.52	3.5178	4.0814	3.63883
0.53	3.6228	4.1454	3.63898

Table 4.3: Spread versus depth at close imaging for square, circular and hexagonal based microlenses.

4.5 Conclusions

A model of the 3D-integral recording camera system, involving both geometrical and physical aspects of the optical process, has been described. Three variants of the camera, one based on square, other based on circular and another based on hexagonal microlenses, have been analysed. The performances of the three systems have been compared in terms of retained angular information and point spread ([57]).

The coordinates of the intermediary, integrated, and recorded data have been computed and used to retrieve an object point position from two recorded micro-intensity distributions. Two types of microlens packaging in the array (rectangular net and honeycomb) have been considered, and the number of recorded micro-intensity distributions for a single physical point has been evaluated for each of the packaging types and in relationship with point's position and depth. The analysis has taken into account the combined effect of all the microlenses of the arrays in the image formation, and has shown that the recording process using honeycomb packed

Point depth $\|z\|$ [m]	Spread ratios at remote imaging		
	$spread_{sq}/spread_{circ}$	$spread_{sq}/spread_{hex}$	$spread_{hex}/spread_{circ}$
0.36	84.68%	93.25%	90.81%
0.37	84.69%	93.26%	90.81%
0.38	84.69%	93.27%	90.80%
0.39	84.70%	93.29%	90.79%
0.40	84.71%	93.31%	90.79%
0.41	84.72%	93.33%	90.77%
0.42	84.73%	93.36%	90.76%
0.43	84.75%	93.39%	90.74%
0.44	84.77%	93.44%	90.72%
0.45	84.80%	93.50%	90.69%
0.46	84.84%	93.59%	90.65%
0.47	84.89%	93.71%	90.59%
0.48	84.97%	93.88%	90.51%
0.49	85.08%	94.14%	90.38%
0.50	85.27%	94.56%	90.17%
0.51	85.59%	95.29%	89.82%
0.52	86.19%	96.67%	89.16%
0.53	87.39%	99.56%	87.78%

Table 4.4: Spread versus depth at remote imaging for square, circular and hexagonal based microlenses. Ratios between spreads corresponding to the types of microlenses considered.

circular or hexagonal microlenses retains more angular details than that corresponding to a rectangularly packed square or circular based microlens array. Moreover, the hexagonal microlens arrays do not have inactive areas and ensure an efficient use of the recording surface ([57]).

The point spread function of the entire process and its dependence on point depth have been analysed for square, circular, and hexagonal based microlens configurations in both close and remote imaging conditions. The use of a camera based on square microlenses is advisable when recording scenes which are reconstructed at a distance away from the recording screen, whilst spherical microlenses are recommended for recording scenes which straddle the recording screen. The hexagonal based microlens system presents a good trade-off between spread performance at both remote and close imaging and the retained angular information. It is therefore recommended when a wide depth range, from very near to very far, is recorded ([57]).

The derivation of the point spread function is of crucial importance for the present research as it shows how depth information about the scene is encoded in the 3D-omnidirectional integral image. The following chapters will make use of the point spread expression to extract depth from integral images.

Chapter 5

Depth Extraction as an Inverse Problem

5.1 Introduction

The mathematical models derived for various configurations of the integral imaging camera system in Chapters 3-4 have shown how depth information is encoded in the planar recording of integral images. It is now possible to newly formulate depth extraction as the inverse problem of image formation and recording. It proves ill-posed in the continuous form and ill-conditioned in the discretised variant, due to the loss of information that occurs at the direct process. In order to cure the ill-posedness / ill-conditioning of the problem, approximate solutions satisfying additional constraints coming from the physics of the problem are searched, using so-called *regularization methods*. The present chapter comprises an extensive overview of existing regularization methods applicable to depth extraction.

5.2 Formulation of depth estimation as an inverse problem

In stereoscopy, depth is traditionally estimated from the disparity of corresponding points recorded in stereo pairs, as explained in Chapter 2. Though integral imaging is very different from stereoscopy, the idea could be useful taking into account that

any object point is recorded behind a number of microlenses and has, therefore, many corresponding intensity distributions in the image plane. The disparity between two corresponding intensity distributions for a given physical point has been derived in Chapters 3-4. The same as in the stereoscopic case, the problem of matching corresponding intensity distributions arises. Unfortunately, due to the resolution achievable in an electronic version of an integral image, which typically is 8-15 pixels per microlens, the disparity yields smaller than the size of a pixel. Therefore, deriving the depth of a point is practically impossible. This is why a novel approach of depth extraction is proposed in the following.

The study of the operation of the integral camera system presented in Chapter 3 and Chapter 4 shows that the spread function of the entire optical process combines the spread effect of both the central double microlens screen and of the recording array ([56], [57]). Its component behind a microlens l which 'sees' the object point (x, y, z) is given by:

$$K_{l(x,y,z)}(X,Y) = \alpha(z)\exp\left(-\frac{(X-X_l)^2}{U_1^2(z)}\right)\exp\left(-\frac{(Y-Y_l)^2}{U_2^2(z)}\right), \tag{5.1}$$

where (X, Y) are the coordinates in the image space, (X_l, Y_l) represents the point of maximum intensity in the microimage of the physical point (x, y, z) formed behind microlens l and $\alpha(z)$, $U_1(z)$, $U_2(z)$ are functions which depend explicitly on lenslet shape and parameters and on point depth z ([56], [57]). All equations refer to the coordinate system depicted in Figure 3.1.

The total point spread function is the sum of the local spreads behind all the microlenses which image the point (x, y, z):

$$K(X,Y,x,y,z) = \sum_{l(x,y,z)} K_{l(x,y,z)}(X,Y)\,. \tag{5.2}$$

Equation (5.1) indicates that the spread function is space variant and that it holds depth information in all the factors. This fact will be used in recovering three dimensional data about the object space.

According to the Rayleigh-Duffieux model ([80], [27], [85]) discussed in Chapter 3, the process of formation of an integral image integrates the intensity f of the

object space with respect to the point spread function. Hence, the image formation equation for the integral imaging system is ([58], [59], [60]):

$$g(X,Y) = \int f(x,y,z) \sum_{l(x,y,z)} K_l(X,Y,x,y,z)dx\,dy\,dz\,, \tag{5.3}$$

where $f(x,y,z)$ is the intensity at an object point (x,y,z).

It is noticed that the integral image formation process *is not a convolution process* because the critical shift invariance property does not hold. Therefore classical deconvolution methods cannot be used to tackle the problem of scene reconstruction from an integral image ([58], [59]).

However, due to the properties of the Riemann-Lebesgue integral ([61], [83]), it is true that the formation of an integral image is a linear process because:

$$\int (a_1 f_1 + a_2 f_2)K = a_1 \int f_1 K + a_2 \int f_2 K \tag{5.4}$$

for any $a_1, a_2 \in R$ and any intensity distributions $f_1, f_2 : R^3 \to R$. Therefore, the point spread function provides a linear image formation operator A such that Equation (5.3) is written as:

$$g = A\,f\,. \tag{5.5}$$

Deriving the intensity f of the object space allows the estimation of the depth of the recorded scene. Namely, the points with null or very small reconstructed intensity are empty, and the object points with high reconstructed intensity are on the visible, thus imaged, part of an object in the scene, and it is possible to draw the *depth map* of the object space from the *intensity map* f ([59]).

In this way, object space intensity estimation and, subsequently, depth extraction from integral images appear as the reverse process of image formation and recording: *given the image g, which is the recorded* **'effect'** *of the luminous intensity in the object space, find the* **'cause'** *that had produced it, namely the intensity f of the scene*([58], [59]). This is a typical **inverse problem**([88], [6]). The direct problem – image formation and recording – was studied in Chapters 3-4 and in [56], [57].

Given this formulation of the problem ([58], [59]), the aim of the present work is to derive the intensity f from Equations 5.3 and 5.5 when the image g and the

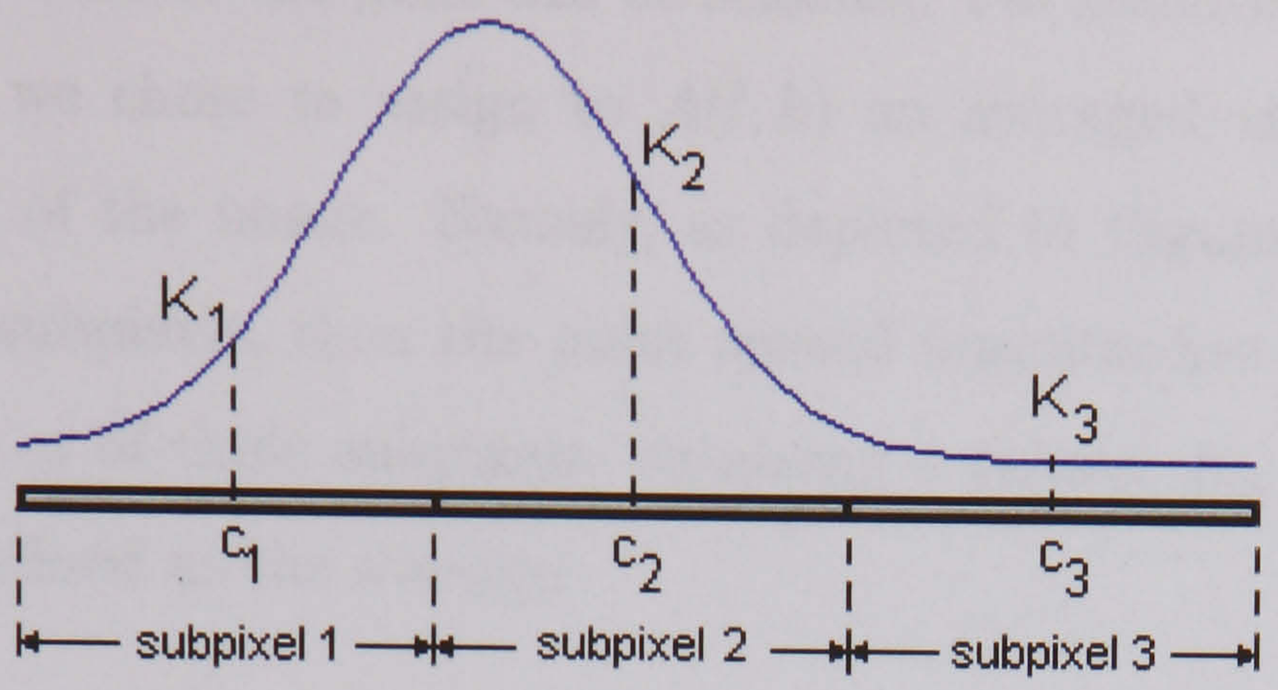


Figure 5.1: Subpixelation used at the computation of the imaging operator.

point spread function K of the optical system are known. It is noticed that this approach of the depth estimation problem allows eventually the retrieval not only of the depth of the scene, but also of *a reconstruction of the object space.*

5.3 Discretisation of the problem

A numerical approach of the problem requires a discretisatisation of both the imaged space and the image. The image has a natural discretised structure, as it is electronically represented by a two dimensional $M \times N$ array of pixels which can be read into a vector $g \in R^m$, $m = M \times N$ being the total number of pixels in the image.

A suitable discretisation of the imaged scene should be also considered. This can be done by conceiving the object space not as a continuous structure, but as a set of n sampling points such that the intensity associated with this sampling structure is a vector $f \in R^n$. The positions of these n points can be arbitrarily chosen such that they cover a relevant part of the object space.

The linear imaging operator A associated with the point spread function for a specific sampling of the object space is a real matrix $m \times n$ whose generic element $A(l,k)$ defines the intensity produced by the point k of the object space sampling, with unit intensity, in pixel l of the image. The value assigned to $A(l,k)$ is computed using the expression of the point spread function K, which is continuous and takes a whole range of values across the surface of the pixel l. The discretisation requires that a sampled value of the point spread function be chosen. For instance, the

value of K in the centre of the pixel can be selected. For smoothness and increased accuracy reasons, we chose to assign to $A(l,k)$ an averaged value corresponding to a subpixelation of the image. Namely, as depicted in Figure 5.1, a pixel l has been divided in s subpixels, then the point spread function has been computed in the centres $c_1, \ldots, c_s$ of these subpixels, obtaining s values: $K_1, \ldots, K_s$, and then $A(l,k)$ has been defined as the average:

$$A(l,k) = \frac{K_1 + K_2 + \cdots + K_s}{s}. \tag{5.6}$$

Generally, subpixelations using $s = 3, 4, 5$ subpixels have been used. Higher subpixelations involve more expensive computations without improving significantly the accuracy of the resulting imaging operator.

In conclusion, given this discretised approach, Equation (5.5) becomes a vectorial equation that can be easily used in numerical applications. The discrete problem thus obtained is well-posed, as proved by the following result:

THEOREM 5.1 (UNIQUENESS OF THE SOLUTION, [59]) *The discrete equation (5.5) has a unique solution in the case of noise free image formation.*

Proof. In the ideal case of noise free imaging conditions, equation (5.5) has a solution $f^* \geq 0$. The solution f^* is assumed sparse as it corresponds to a wide space with relatively few radiating points. Supposing that there is a $\Delta f \in R^n$ such that $A\Delta f = 0$, then it is natural to assume that Δf has frequent interleavings +/-, due to the fact that A has only positive elements. These interleavings are very likely to correspond to positions where the elements of the vector f^* are null. Therefore, $f^* + \Delta f$ has both positive and negative components, thus it cannot be a (positive) solution of the imaging equation (5.5). It follows that f^* is the unique solution of (5.5). ■

As already discussed in Chapter 3, the two tier camera system produces a 3D-optical model of the object space, which is then captured by the recording device. Obtaining the configuration of the optical model is equivalent to obtaining the configuration of the object space, as they are symmetric with respect to the central plane of the two tier camera system (Figure 5.2). For simplicity reasons, but without any

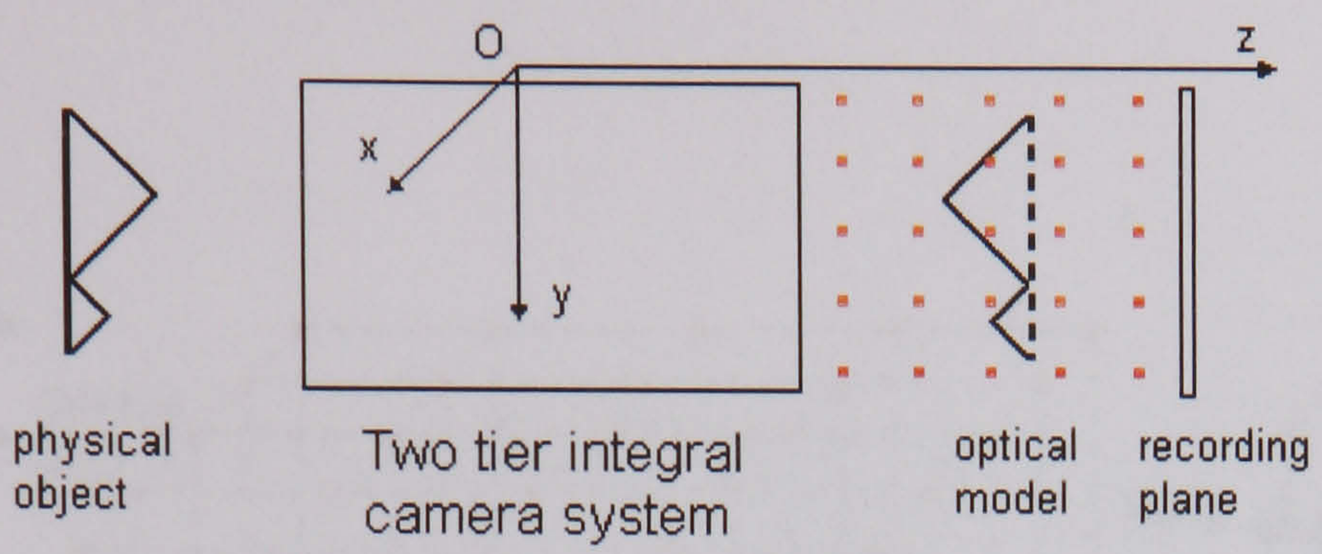


Figure 5.2: The position of a sampling net with respect to the integral camera system (upper view). The portion of space where the optical model lies is sampled.

loss of generality, throughout this work, the physical space where the optical model lies is sampled and the optical model of the object space is reconstructed. The situation is depicted in Figure 5.2. In the following, 'sampling the object space' means 'sampling the space where the optical model is formed'. The phrase is justified by the fact that the optical model of the object is the actual recorded object.

5.3.1 Particular case: discretisation of the problem for the unidirectional integral imaging system

The reader is reminded that the unidirectional (lenticular) integral imaging system uses cylindrical surface microlenses in the recording array to produce integral images which exhibit horizontal parallax. The 2D-cylindrical refraction model already discussed in Chapter 3 is a very good approximation of the reality ([12]), and, according to this model, physical points in a plane $x = x_0$ (Figure 3.1, Figure 5.3) are recorded on the line $g(., j_0)$ of the image situated in the plane $x = x_0$. Moreover, according to this model, the imaging operator $A_{\{x=x_0\}}$ computed for a set of sampling points $\{Q_k(x_0, y_k, z_k),\ k = 1, \ldots, n\}$, situated in a plane $x = x_0$ is the same as the imaging matrix $A_{\{x=x_1\}}$ computed for the corresponding set $\{Q'_k(x_1, y_k, z_k),\ k = 1, \ldots, n\}$, in any other parallel plane $x = x_1$:

$$A_{\{x=x_0\}} \equiv A_{\{x=x_1\}} \equiv A\,. \tag{5.7}$$

Suppose that $g(s,t)$, $1 \leq s \leq M$, $1 \leq t \leq N$, is a pixelated lenticular integral image represented as a matrix, and that the object space is sampled with a three

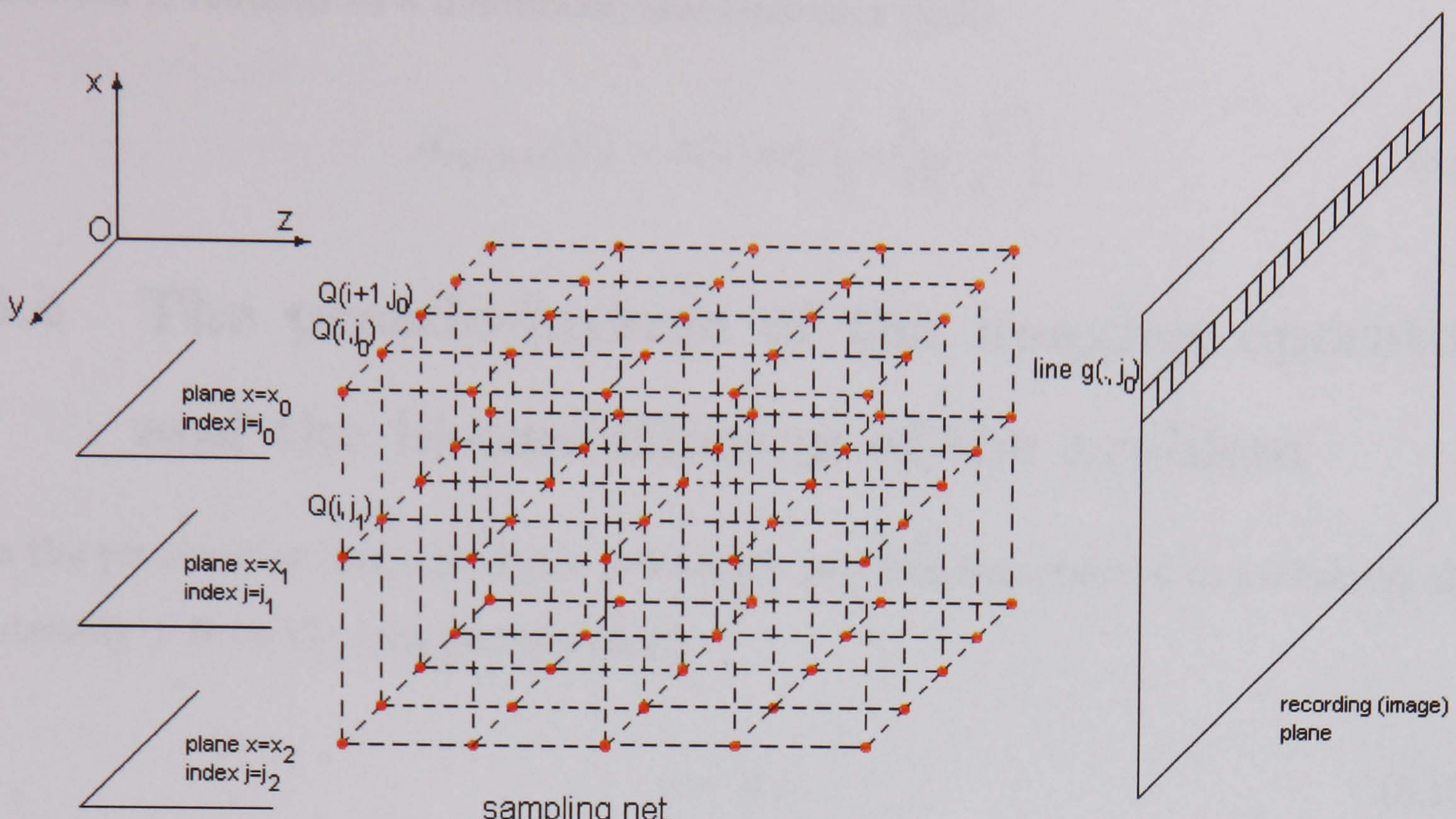


Figure 5.3: Sampling and unidirectional integral imaging.

dimensional net of points $Q(i,j)$, $1 \leq i \leq n$, $1 \leq j \leq N$, placed in parallel planes, as shown in Figure 5.3. i is the index of a point in an $x = const$ plane, and j is the index of the x-plane. If $f(i,j)$ denotes the intensity in the point $Q(i,j)$, then, because of the identity (5.7), it holds true that:

$$g(.,j_0) = A\, f(.,j_0)\,, \tag{5.8}$$

for any $j_0 \in \{1,\ldots,N\}$. It follows that, in the particular case of unidirectional integral imaging, each line $g(.,j_0)$ of the image can be independently used to reconstruct the intensity distribution in a certain plane $x = x_0$ of the object space. This is a very important feature of the unidirectional imaging system, as it decreases the computational cost of the inversion problem.

Another simplification in the case of unidirectional integral imaging is possible at the computation of the imaging operator A. Due to the small values of the parameters involved in the point spread function equation, only the significant factors should be considered in a computational approach of the depth extraction problem. When using lenticular arrays, i.e. cylindrical surface microlenses, the spread on the x-dimension is up to 100 times smaller than the spread on the y-direction, so the x-factors in the point spread expression can be dropped. Therefore, the point spread

function is reduced to a unidimensional Gaussian ([59]):

$$K_{l(x,y,z)}(Y) = \alpha(z) \exp\left(-\frac{Y - Y_l}{U_2^2(z)}\right) . \tag{5.9}$$

5.4 The pseudo-inverse of the imaging operator and the ill-conditioning of the problem

In the previous section, the depth estimation problem was reduced to evaluating the intensity f from the imaging equation:

$$g = A\, f\,. \tag{5.10}$$

The inverse transformation corresponding to the imaging operator A is represented either by the inverse matrix A^{-1}, if A is nonsingular, or by the Penrose-Moore pseudo-inverse A^{+} ([10], [17], [4]), when A is singular. Multiplying with the pseudo-inverse of A, the reconstruction of the object space is obtained as:

$$f^{+} = A^{+} g\,. \tag{5.11}$$

The existence of the pseudo-inverse for an $m \times n$ operator A with $m \geq n$ is insured by the Singular Value Decomposition theorem:

THEOREM 5.2 (SVD, [10], [5]) *For any matrix A with m rows and n columns, $m \geq n$, of rank p, there exists an orthogonal $m \times p$ matrix U, a diagonal $p \times p$ matrix $\Sigma = diag(\sigma_i)$, and an orthogonal $n \times p$ matrix V such that $A = U\Sigma V^T$. $\sigma_1 \geq \ldots \geq \sigma_p > 0$ are called the* singular values *of the matrix A.*

If the matrix A is square ($m = n$) and non-singular, then its pseudo-inverse coincides with the inverse and is given by: $A^{+} = A^{-1} = V \cdot diag(1/\sigma_i) \cdot U^T$ ([10], [17]). If the matrix is singular or non-square, then ([10], [17])

$$A^{+} = V \cdot \begin{pmatrix} diag(1/\sigma_i)_{1 \leq i \leq p} & 0 \\ 0 & 0 \end{pmatrix} \cdot U^T , \tag{5.12}$$

where p is the rank of A.

The condition number of the operator A is defined as the ratio between the maximum and the minimum singular values ([5]):

$$cond(A) = \frac{\sigma_1}{\sigma_p}. \tag{5.13}$$

The condition number is very important in amplifying the errors in the initial conditions ([5]). Suppose that the operator A is perturbed with a very small relative error δA such that $\|\delta A\| \cdot \|A^+\| \ll 1$. Suppose also that the image vector g is affected by an error or noise δg. Then, an error δf will necessarily occur in the solution f such that $(A+\delta A)(f+\delta f) = g+\delta g$. An estimation of δf from this equation yields ([5]):

$$\frac{\|\delta f\|}{\|f\|} \leq \frac{cond(A)}{1 - cond(A)\|\delta A\|/\|A\|} \left(\frac{\|\delta g\|}{\|g\|} + \frac{\|\delta A\|}{\|A\|} \right), \tag{5.14}$$

where the equality can be reached. It is noticed that a high condition number is related to an amplification of the initial perturbation of the data which results in the instability of the solution of Equation (5.5) because $\|\delta A\|/\|A\|$ is very small and thus the amplification can be proportional with the condition number.

Unfortunately, the condition number of the integral imaging operator A is very high, which makes the solution f^+ very unstable ([5]), i.e. very sensitive to inherent perturbations of the image g and of the operator A. A typical perturbation is the estimation itself of the imaging operator A. It is computed using a number of samplings less or equal to the number of pixels in the image, and thus only a rough approximation, very ill-conditioned, of the continuous operator A can be obtained. This initial error is amplified in the result obtained by pseudo-inversion such that the solution f^+ is not physically meaningful.

This assertion is confirmed by the following example: the noiseless computer generated image of a cube was considered. The object space was sampled and the imaging operator associated to this sampling was computed. Then the generalised inverse $f^+ = A^+ g$ was computed. A transversal section of this reconstruction is presented in Figure 5.4, together with the actual contour of the object (the two visible faces of the cube). It is noticed that f^+ gives no useful indication about the shape or position of the object. Moreover, the intensity values obtained for f^+ vary

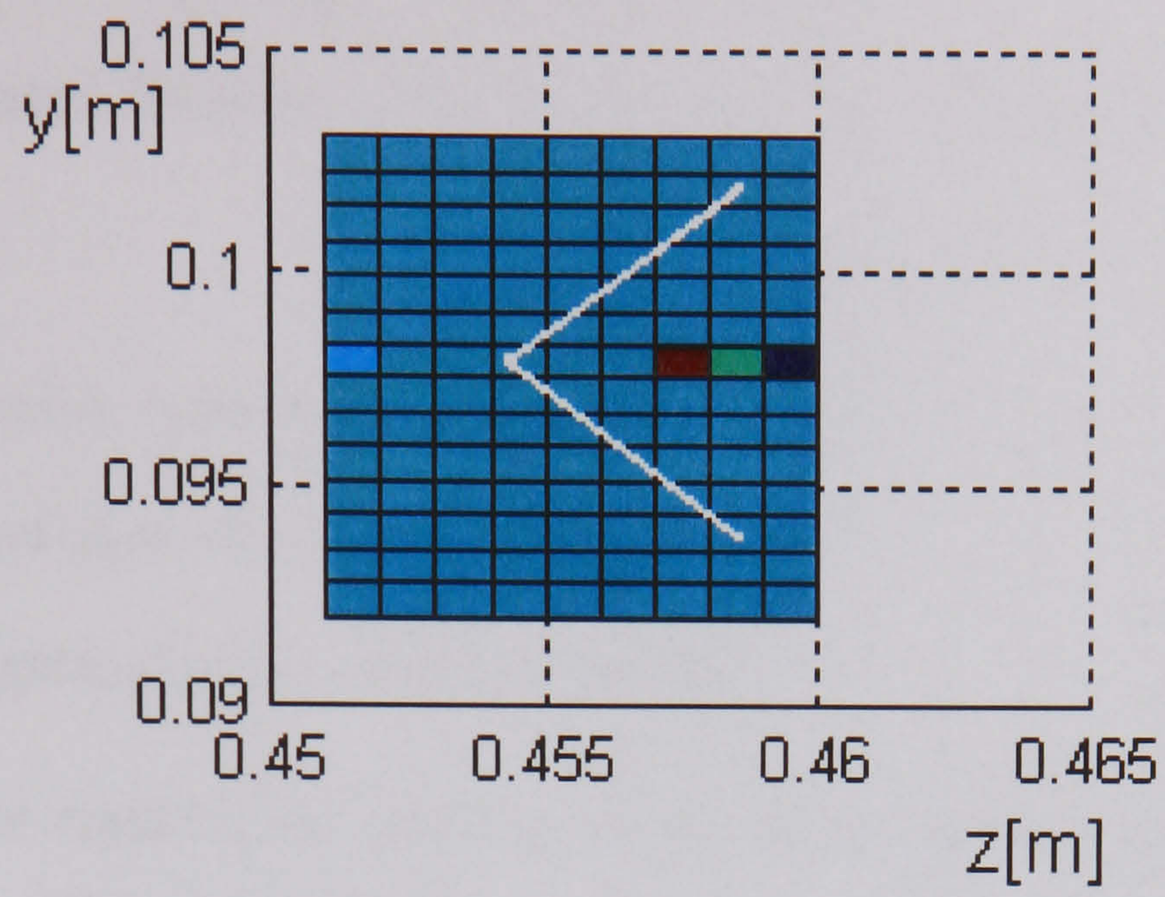


Figure 5.4: The reconstruction of the visible faces of an opaque cube from a noiseless integral image using the pseudo-inverse of the imaging operator.

between -10^5 and 10^6, which is physically meaningless for the given object.

As the example has shown, the vector $f^+ = A^+ g$ is usually a meaningless bad approximation of the real object space intensity distribution f. This is a situation where the parameters in the state vector f are discretised values of a function at points of a suitable grid. Refining the coarse grid increases the number of parameters, and, because the amount of grid independent data is finite (g), a very fine grid cannot be used with standard inversion methods which require the number of samplings to be lower or equal to the number of data.

To obtain sensible parameter estimates in such a situation, it is necessary to use supplementary qualitative information coming from the physics of the problem. So-called *regularization methods*, which are families of operators that approximate A^+ pointwise in a suitable way, are needed to obtain physically realistic estimates for the solution of this ill-conditioned problem.

5.5 Overview of the regularization methods applied to depth extraction

It is beyond the scope of this thesis to review all the existing regularization methods. The methods considered for application to depth extraction can be classified as:

- non-iterative:

1. Tikhonov regularization ([88], [89], [90])

2. Truncated Singular Value Decomposition (TSVD) ([6])

- iterative:

1. Landweber type methods ([48])

2. Steepest descent method ([6],[41])

3. Conjugate gradient method ([6],[42])

In general, *the regularizing property is proved by showing convergence for exact data and stability for perturbed data.*

The continuous problem (5.3) underlying the depth extraction discretised problem is a Fredholm integral equation of the first kind. The Fredholm integral equation has been approached under the frame of the functional analysis ([88], [1]), and regularization methods and convergence criteria have been proposed for the infinite-dimensional problem ([88],[31],[96],[18]). Recently, some of these theoretical approaches have been translated into linear algebra language ([71]), as an application to discrete data, and used in image deconvolution and restoration ([6], [7], [8]). Depth extraction is *not a deconvolution problem*, but the existing regularization methods, in their most general form, can be adapted and used to obtain object space reconstructions from integral images.

The present chapter summarises the most important regularization methods suitable for application to the current problem and their convergence properties.

5.5.1 Unconstrained and constrained least-squares solutions

The object space reconstruction from planar recorded data of 3D-integral images is a linear inverse problem that requires to solve the equation:

$$A f = g\,, \tag{5.15}$$

where A is the linear and continuous *imaging operator*, g is the *recorded image*, i.e. the datum of the problem, and f is the *unknown intensity map of the object scene.* The image g is affected by noise or experimental / scanning errors, hence it has the

structure ([6]):

$$g = A\, f^{(0)} + \nu\,, \tag{5.16}$$

where $f^{(0)}$ is the function to be estimated and ν is a term describing the noise.

The following notations are used:

- $\mathcal{R}(A)$ = the range of the operator A.
- $\mathcal{N}(A)$ = the null-space of the operator A.
- $A(\mathcal{C})$ = the image of the set $\mathcal{C}$ through the operator A.
- A^* = the adjoint operator of A.
- The singular value decomposition (SVD) of the $m \times n$ linear operator A of rank p: $A = U\Sigma V^T$, where U and V are orthogonal matrices of dimensions respectively $m \times p$ and $n \times p$, and Σ is the diagonal matrix defined by the singular values of A: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p$.

If the range of A, $\mathcal{R}(A)$, is not closed, then the problem (5.15) generally has no solution when g has the structure (5.16) because ν does not belong to $\mathcal{R}(A)$, and the problem is said to be *ill-posed* ([6]). Therefore, in the search for a meaningful solution the least-squares approach is adopted. An (unconstrained) **least-squares solution** is a function f which satisfies:

$$\|A\, f - g\| = min\,. \tag{5.17}$$

The non-constrained least-squares problem (5.17) is still ill-posed if the range $\mathcal{R}(A)$ is not closed, and least-squares solutions exist if and only if $g \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$ ([79]). The set of least-squares solutions coincides with the set of solutions of the Euler equation ([6]):

$$A^*\, A\, f = A^*\, g\,. \tag{5.18}$$

The minimal norm solution is the *generalised solution* f^+, which is also the unique least-squares solution belonging to $\mathcal{N}(A)^{\perp}$ ([6]). The *generalised inverse* A^+ is the linear operator with domain $\mathcal{D}(A^+) = \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$ and range $\mathcal{R}(A^+) = \mathcal{N}(A)^{\perp}$

([10]) which satisfies:

$$f^{+} = A^{+} g \,. \tag{5.19}$$

However, if $f^{(0)}$ from equation (5.16) belongs to some closed and convex set $\mathcal{C}$, the use of linear regularization methods does not provide approximate solutions which belong to $\mathcal{C}$ ([6]). In this case, the initial problem (5.15) can be reformulated as a **constrained least-squares problem**:

$$\|A\, f - g\| = min\,, \quad f \in \mathcal{C} \tag{5.20}$$

The constrained problem can still be ill-posed if $\mathcal{R}(A)$ is not closed, and existence of solutions depends on the set $\mathcal{C}$ ([79]). If $A(\mathcal{C})$ denotes the image of $\mathcal{C}$, which is also a convex set, the following cases can be distinguished:

1. The problem (5.20) has a solution for any image g if and only if the set $A(\mathcal{C})$ is closed. This condition is satisfied, for instance, when $\mathcal{C}$ is bounded ([79]).

2. If $A(\mathcal{C})$ is not closed, then the problem (5.20) has a solution if and only if the convex projection of g onto $\overline{A(\mathcal{C})}$ belongs to $A(\mathcal{C})$. For instance, if $\mathcal{C}$ is the set of non-negative functions in $L^2(R)$, then $A(\mathcal{C})$ is not closed ([79]).

Depth extraction from 3D-integral images is approached as a constrained least-squares problem, and least-squares solutions of the imaging equation which are positive and upper bounded are sought. Therefore, when using the constrained regularization methods, the closed convex set $\mathcal{C}$ is:

$$\mathcal{C} = \left\{ f \in L^2(R) | 0 \leq f \leq T \right\} , \tag{5.21}$$

or, for numerical applications, the discretised version of (5.21). T is a positive upper threshold that is not necessarily finite. Generally, the set $A(\mathcal{C})$ is not closed.

Regularization methods have a very wide range of application in Mathematics and Physics. A number of them are reviewed in the following sections, the presentation being biased towards the current problem. Most equations can be interpreted both in terms of functional analysis (continuous linear operators and L^2-functions) or linear algebra (matrices and vectors). Due to the fact that the numerical appli-

cation to integral imaging is performed on discrete data, the linear algebra language has been preferred. However, the main results are true in the general frame of continuous linear operators.

5.5.2 Tikhonov Regularization

Unconstrained Tikhonov regularization

Among the methods developed to solve linear ill-posed inverse problems, Tikhonov regularization is the most popular. In this method, a family of regularized solutions f_μ of the problem (5.17) is defined as the family of functions which minimise the following functional ([88], [6]):

$$\Phi_\mu(f,g) = \|Af - g\|^2 + \mu\|f\|^2 . \tag{5.22}$$

μ is called *regularization parameter*. It can be proved that a function f_μ is a minimal point of the Tikhonov functional $\Phi_\mu(f,g)$ if and only if it is a solution of the Euler equation ([88], [6]):

$$(A^*A + \mu I)f = A^*g . \tag{5.23}$$

Therefore, the problem is reduced to obtaining the solutions of the equation (5.23). The solution has the form:

$$f_\mu = \Psi_\mu g , \tag{5.24}$$

where the operator Ψ_μ is given by:

$$\Psi_\mu = (A^*A + \mu I)^{-1}A^* . \tag{5.25}$$

In order to prove the existence of the solution to the Euler equation, an arbitrary element f is decomposed along the singular vectors of the linear operator A and along the null space of A. Using the decomposition $A = U\Sigma V^T$, f can be written as $f = \sum_{j=1}^{p} < f, v_j > v_j + v$, where v_j are the columns of the orthogonal matrix V and v is the 'invisible' component (i.e. $Av = 0$). Equation (5.23) can now be written as:

$$\sum_{j=1}^{p}(\sigma_j^2 + \mu) < f, v_j > v_j + \mu v = \sum_{j=1}^{p} \sigma_j < g, u_j > v_j , \tag{5.26}$$

and there exists a unique solution of minimal norm f_μ, obtained for $v = 0$, given by ([6]):

$$f_\mu = \sum_{j=1}^{p} \frac{\sigma_j}{\sigma_j^2 + \mu} < g, u_j > v_j \,, \tag{5.27}$$

for any noisy image g.

The linear operator Ψ_μ represents an approximation of the generalised inverse A^+ in the sense that:

$$\lim_{\mu \to 0} \Psi_\mu g = A^+ g \,. \tag{5.28}$$

Therefore, the method provides a least-squares solution for each $\mu > 0$. The issue which arises is to choose the solution that suits best the problem being solved. Consequently, the problem is to establish the existence of an optimal value of the regularization parameter μ and, eventually, find a way of determining it for practical applications. In order to further analyse the method from this point of view, the energy functional and the discrepancy functional are necessary.

The *energy functional* is given by ([6]):

$$E^2(f_\mu) = \|f_\mu\|^2 = \sum_{j=1}^{p} \frac{\sigma_j^2}{\sigma_j^2 + \mu} | < g, u_j > |^2 \tag{5.29}$$

and is a decreasing function of the regularization parameter μ.

The *discrepancy functional* ([6]):

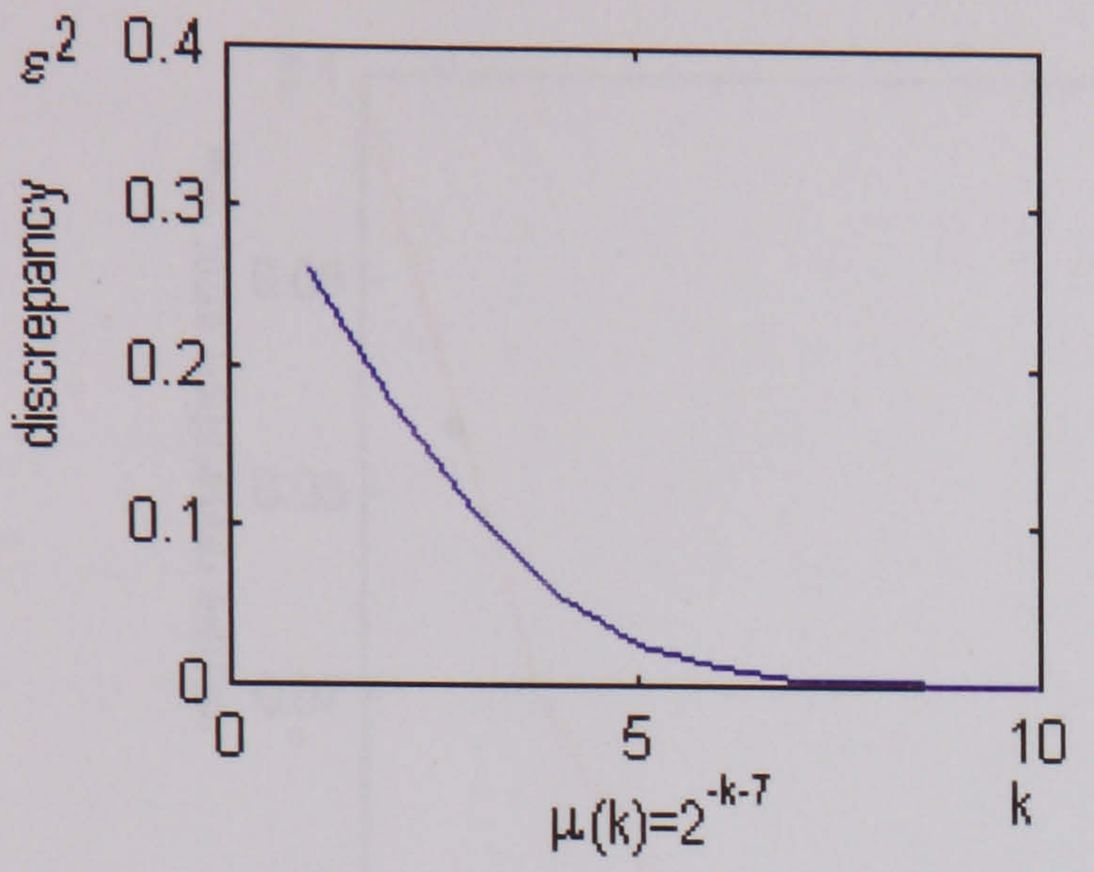
$$\epsilon(f_\mu, g) = \|Af_\mu - g\|^2 = \sum_{j=1}^{p} \frac{\mu^2}{\sigma_j^2 + \mu} | < g, u_j > |^2 + \|u\|^2 \tag{5.30}$$

is an increasing function of the regularization parameter μ.

Figure 5.5 depicts the variation of the discrepancy and energy functionals when the regularization parameter μ converges to 0.

The difference between the regularized object f_μ obtained from a noisy image g and the true object $f^{(0)}$ is therefore:

$$f_\mu - f^{(0)} = (\Psi_\mu A f^{(0)} - f^{(0)}) + \Psi_\mu \nu \,. \tag{5.31}$$



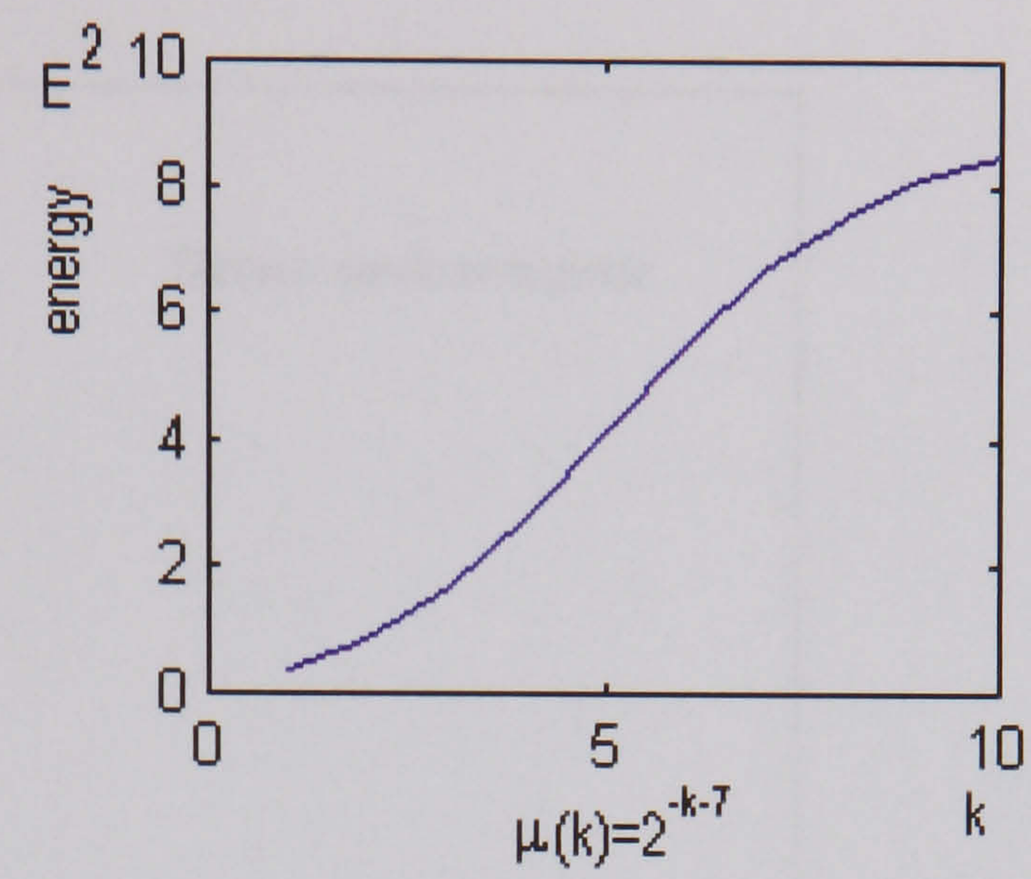
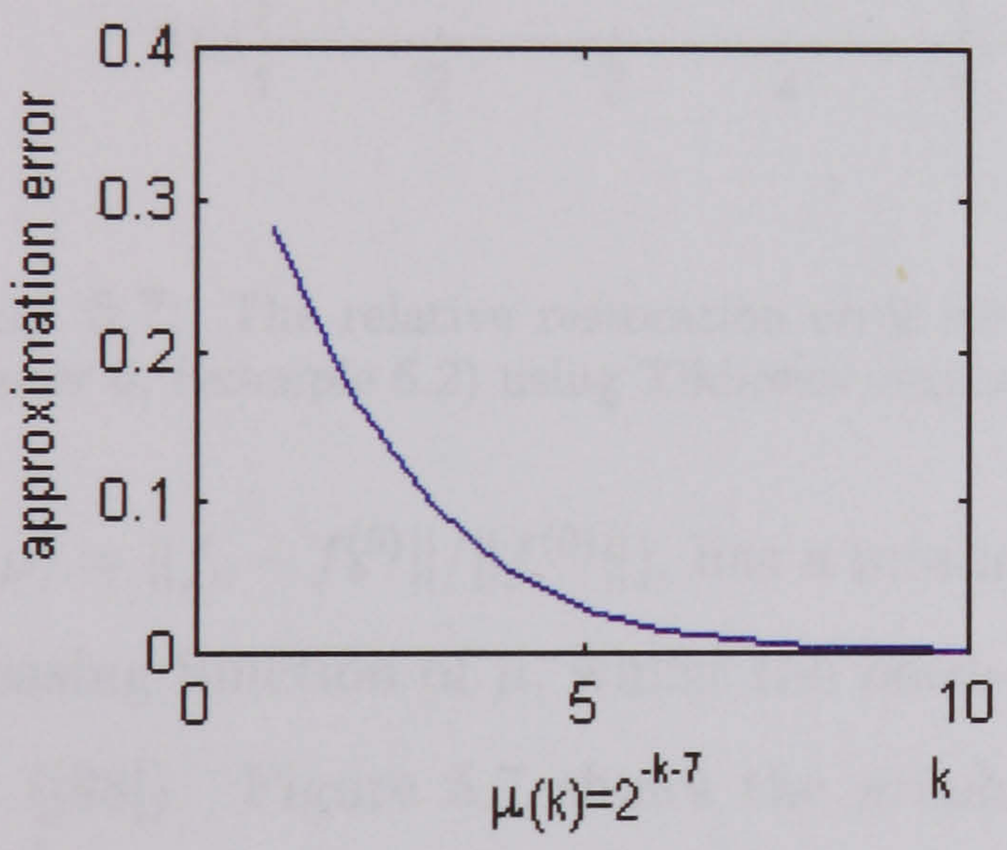


Figure 5.5: The discrepancy (a) and energy (b) as functions of the regularization parameter for the reconstruction of a nine point object space (Chapter 6, Example 6.2) using Tikhonov regularization.



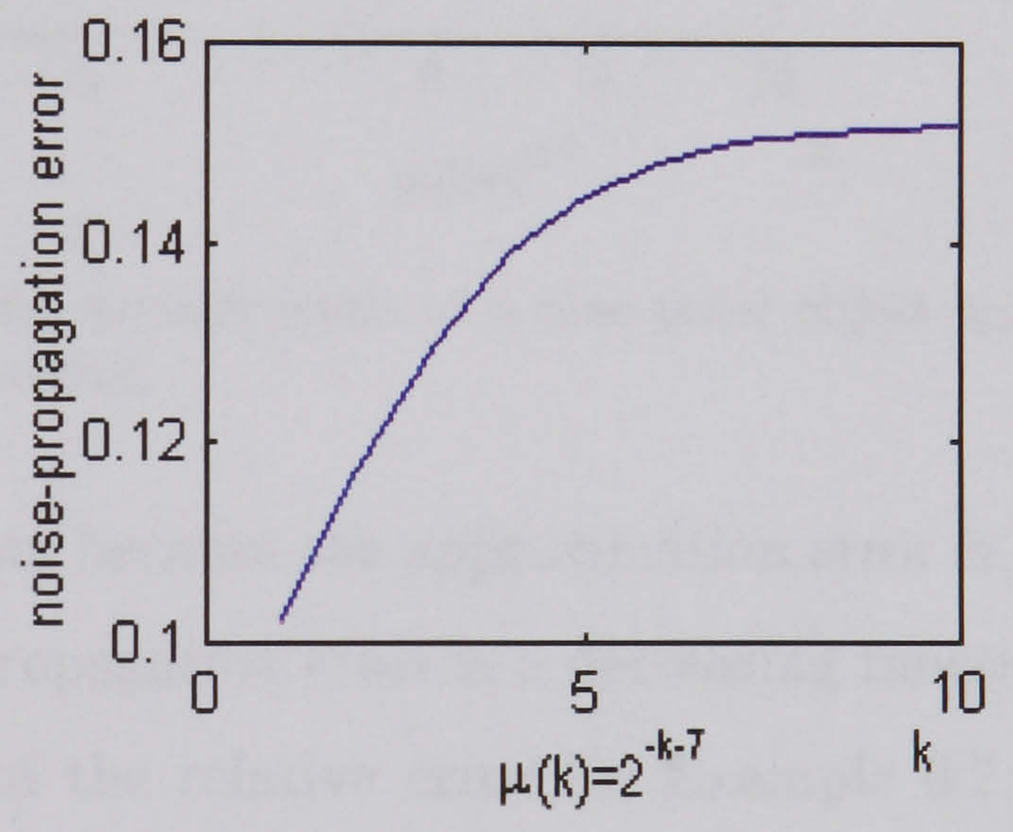


Figure 5.6: The approximation error and the noise propagation error for the reconstruction of a nine point object space (Chapter 6, Example 6.2) using Tikhonov regularization.

It follows that the *approximation error* is ([6]):

$$\|\Psi_\mu A f^{(0)} - f^{(0)}\|^2 = \sum_{j=1}^{p} \frac{\mu^2}{(\sigma_j^2 + \mu)^2} | < f^{(0)}, v_j > |^2 + \|v^{(0)}\|^2 , \qquad (5.32)$$

where $v^{(0)}$ is the 'invisible' component of $f^{(0)}$, and the *noise-propagation error* is ([6]):

$$\|\Psi_\mu \nu\|^2 = \sum_{j=1}^{p} \frac{\sigma_j^2}{(\sigma_j^2 + \mu)^2} | < \nu, u_j > |^2 . \qquad (5.33)$$

Figure 5.6 depicts the variation of the approximation error and noise propagation error when the regularization parameter μ converges to 0.

The *total restoration error*, $\rho(\mu) = \|f_\mu - f^{(0)}\|$ (or the relative restoration error

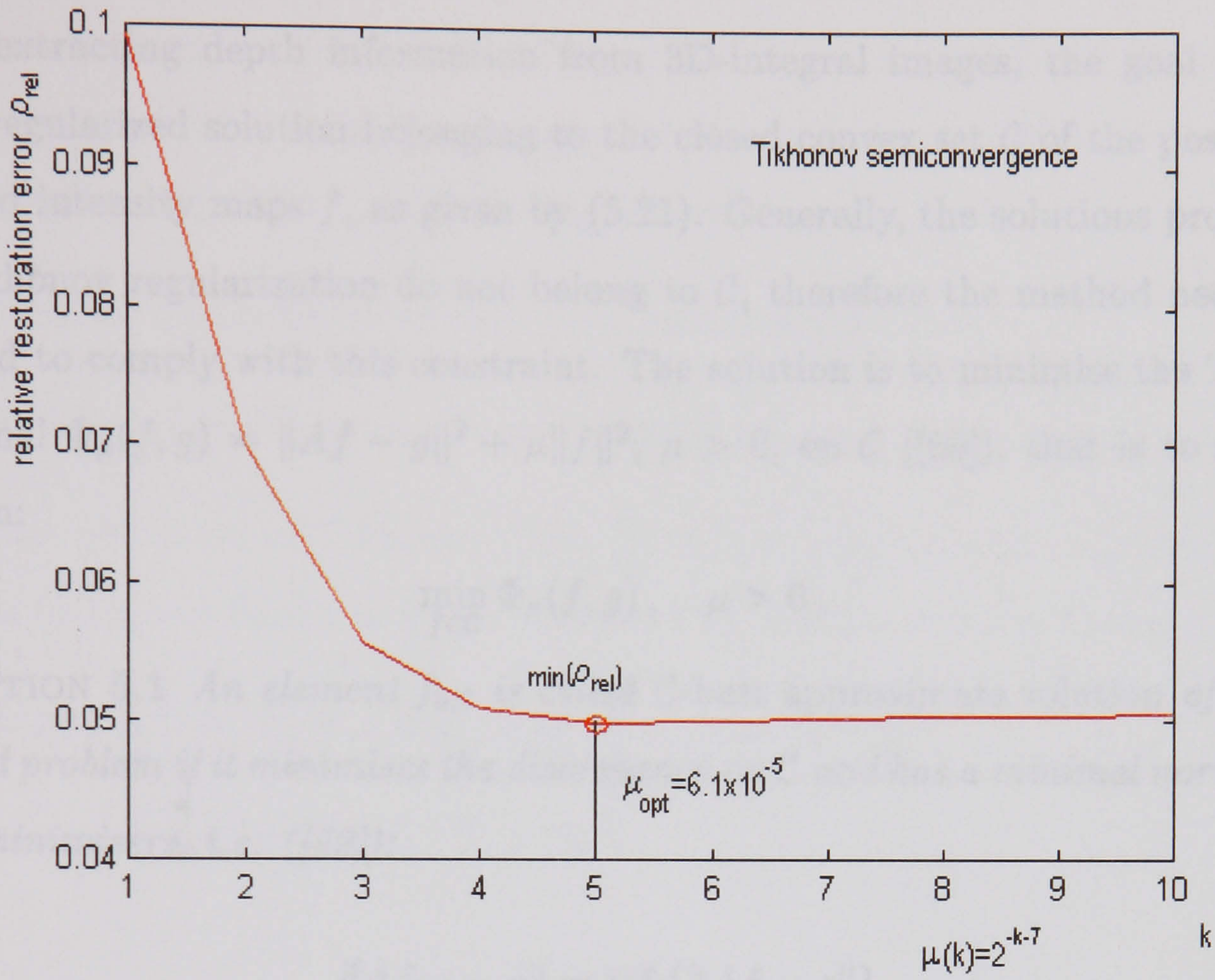


Figure 5.7: The relative restoration error for the reconstruction of a nine point object space (Chapter 6, Example 6.2) using Tikhonov regularization.

$\rho_{rel}(\mu) = \|f_\mu - f^{(0)}\|/\|f^{(0)}\|$), has a minimum because the approximation error is an increasing function of μ, whilst the noise-propagation error is a decreasing function of μ ([88]). Figure 5.7 shows the graph of the relative error for Example 6.2 in Section 6. Tikhonov regularization has therefore a **semiconvergence property** in the case of noisy images and it is possible to find the value μ_{opt} of the regularization parameter which provides the best reconstruction of the object ([88]).

The practical determination of the optimal regularization parameter is not a simple problem. It can be approached using the discrepancy principle of Morozov ([67]), the energy principle ([6]), cross validation or Miller's method ([65], [6]). The most important *rules* which govern the choice of the optimal parameter for most regularization methods are summarised at the end of the chapter. A number of algorithms for fast realization of the parameter choice in the case of the Tikhonov method and other inverse problems have been reported ([30], [97]).

Constrained Tikhonov regularization

When extracting depth information from 3D-integral images, the goal is to obtain a regularized solution belonging to the closed convex set $\mathcal{C}$ of the positive and bounded intensity maps f, as given by (5.21). Generally, the solutions provided by the Tikhonov regularization do not belong to $\mathcal{C}$, therefore the method needs to be adjusted to comply with this constraint. The solution is to minimise the Tikhonov functional $\Phi_\mu(f,g) = \|Af - g\|^2 + \mu\|f\|^2$, $\mu > 0$, on $\mathcal{C}$ ([69]), that is to solve the problem:

$$\min_{f\in\mathcal{C}} \Phi_\mu(f,g)\,, \quad \mu > 0\,. \tag{5.34}$$

DEFINITION 5.1 *An element $f_{0,\mathcal{C}}$ is called* $\mathcal{C}$-best approximate solution *of the constrained problem if it minimises the discrepancy on $\mathcal{C}$ and has a minimal norm among other minimisers, i.e. ([69]):*

$$\|Af_{0,c} - g\| = \inf_{f\in\mathcal{C}} \{\|Af - g\|\}\,, \tag{5.35}$$

and

$$\|f_{0,c}\| = \inf_{f\in\mathcal{C}} \{\|f\| \,|\, \|Af - g\| = \ |Af_{0,c} - g\|\}\,. \tag{5.36}$$

It can be proved ([69]) that a $\mathcal{C}$-best approximate solution exists if and only if the projection $P_{\overline{A(\mathcal{C})}}g \in A(\mathcal{C})$; it is then unique. Neubauer ([69]) has shown that the constrained problem (5.34) has a unique solution for any $\mu > 0$ and that these solutions converge to the $\mathcal{C}$-best approximate solution (which is unique) for $\mu \to 0$ if $P_{\overline{A(\mathcal{C})}}g \in A(\mathcal{C})$. Therefore, the following theorem holds:

THEOREM 5.3 (EXISTENCE AND UNIQUENESS OF THE SOLUTION, [69]) *1. The constrained Tikhonov-regularized solutions $f_{\mu,\mathcal{C}}$ converge to an element in $\mathcal{C}$ for $\mu \to 0$ if and only if $P_{\overline{A(\mathcal{C})}}g \in A(\mathcal{C})$.*

2. $P_{\overline{A(\mathcal{C})}}g \in A(\mathcal{C})$ implies that $\lim_{\mu\to 0} f_{\mu,\mathcal{C}} = f_{0,\mathcal{C}}$.

Moreover, for all $\mu > 0$, the solution depends continuously on the data g ([69]), so the constrained problem is well-posed. The constrained solution $f_{\mu,\mathcal{C}}$ is characterised via the Kuhn-Tucker theory ([55]) as the unique element in $\mathcal{C}$ such that:

$$< A^*Af_{\mu,\mathcal{C}} + \alpha f_{\mu,\mathcal{C}} - A^*g, h - f_{\mu,\mathcal{C}} > \ \geq 0\,, \quad for \quad any \quad h \in \mathcal{C}\,. \tag{5.37}$$

Generalised Tikhonov type algorithms

The Tikhonov regularization method presented in Section 5.5.2 can be regarded as a particular case of a large class of algorithms. In order to define the notion of a 'regularization algorithm', the following assumptions are made:

1. Equation (5.15) has a solution (not necessarily unique).
2. Well-posed problems can be solved accurately.

By the definition of Tikhonov ([89], [2], [6]), a 'regularization algorithm' of the problem (5.15) is a family of linear operators Ψ_μ which has the following properties:

1. for any $\mu > 0$, Ψ_μ is a linear and continuous operator.
2. for any noise-free image $g^{(0)} = Af^{(0)}$:

$$\lim_{\mu \to 0} \Psi_\mu g^{(0)} = A^+ g^{(0)} \tag{5.38}$$

where the limit is in the L^2-norm.

In equation (5.38), $f_\mu = \Psi_\mu g$ is the family of *regularized solutions* and μ is the *regularization parameter.*

Such a class of regularization algorithms in the sense of Tikhonov has been presented by Bakushinskii ([2]).

Bakushinskii considers a bounded function $\phi(\lambda, \mu)$ with real values, defined for all the eigenvalues λ in the spectrum $S(A^*A)$ of the self-adjoint operator A^*A, which satisfies the following conditions:

$$\begin{aligned} \sup_{\lambda \in S(A^*A)} \frac{|\phi(\lambda,\mu)|}{\sqrt{\lambda}} &= K_\mu < \infty \quad (\mu \neq 0) \\ \phi(0,\mu) &= 0\,, \quad \mu \neq 0 \\ \lim_{\mu \to 0} \phi(\lambda,\mu) &= 1 \quad \lambda \in S(A^*A), \lambda \neq 0 \end{aligned} \tag{5.39}$$

The following function is then defined ([2]):

$$\psi(\lambda,\mu) = \begin{cases} \dfrac{\phi(\lambda,\mu)}{\lambda}\,, & \lambda \neq 0 \\ L, & |L| < \infty\,,\ \lambda = 0 \end{cases} \tag{5.40}$$

Using this function, there exists a self-adjoint operator $\Psi_\mu(A^*A)$ defined by ([2]):

$$\Psi_\mu(A^*A)h = \sum_j \psi(\lambda_j, \mu) < h, v_j > v_j \tag{5.41}$$

for any function h. The operator $\Psi_\mu(A^*A)A^*$ is bounded and the procedure for finding:

$$f_\mu = \Psi_\mu(A^*A)A^*g \tag{5.42}$$

is a regularization algorithm of the equation (5.15) in the sense of Tikhonov ([2]).

The Tikhonov regularization presented in 5.5.2 is obtained from this latter algorithm for $\psi(\lambda, \mu) = (\lambda + \mu)^{-1}$ observing that the eigenvalues of the operator A^*A are the same as the squares of the singular values of the operator A: $\lambda_j = \sigma_j^2$.

Two other examples of regularization algorithms ([2]) can be obtained taking:

$$\psi(\lambda, \mu) = \begin{cases} 1/\lambda\,, & \lambda > \mu \\ 0\,, & 0 \leq \lambda \leq \mu \end{cases} \tag{5.43}$$

that results in the removal from the solution of the components which correspond to small singular values (i.e. Truncated SVD method), or:

$$\psi(\lambda, \mu) = \begin{cases} \dfrac{1 - (1 - \tau\lambda)^{1/\tau}}{\lambda}\,, & \lambda \neq 0 \\ \dfrac{\tau}{\mu}\,, & \lambda = 0\,, \end{cases} \tag{5.44}$$

with $0 \leq \tau \leq 2/\|A^*A\|$.

Tikhonov type regularization methods have been also derived for nonlinear inverse problems ([70], [37], [54], [87]), but they exceed the scope and the goal of the present work, which is well encompassed by the frame of a linear problem. Convergence rate results for Tikhonov regularization can be found in [16].

5.5.3 Truncated Singular Value Decomposition

Unconstrained TSVD

Using the singular value decomposition (SVD) of the linear operator $A = U\Sigma V^T$, the generalised solution of the unconstrained problem (5.15) can be written as ([6]):

$$f = \sum \frac{1}{\sigma_j} < g, u_j > v_j \,. \tag{5.45}$$

Assumption: the singular values are ordered decreasingly: $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p$.

In the case of a noise-free g, $< g, u_j >= \sigma_j < f, v_j >$ is very small for j very large as it corresponds to the component of f in the direction v_j multiplied by σ_j. In the case of a noisy image g, $< g, u_j >$ becomes much larger, reaching the level of the noise in the direction u_j, and consequently any information it contains is swamped by the noise. In either case, the corresponding term in the decomposition of f can be filtered out and assimilated to a component in the null-space of the operator A (i.e. an 'invisible' component) ([6]). The fundamental idea of the Truncated Singular Value Decomposition (TSVD) regularization method ([6],[95]) is to remove all the terms corresponding to singular values smaller than a certain threshold μ. Given the indexation in decreasing order of the singular values, the method involves using only the first J singular values, which are greater than μ, and considering the approximate solution:

$$f_J = \sum_{j=1}^{J} \frac{1}{\sigma_j} < g, u_j > v_j \,. \tag{5.46}$$

There exists a linear operator Ψ_J such that (5.46) be written as:

$$f_J = \Psi_J g \,. \tag{5.47}$$

The TSVD method is a regularization algorithm in the sense of Tikhonov because Ψ_J is a bounded operator that converges to A^+ when $\mu \to 0$ ([6]). Indeed, $J = J(\mu)$ increases to p when μ decreases to 0, which means that $\Psi_J g$ converges to $A^+ g$ when $J \to p$. Moreover, in the case of a noisy g, the method has the semiconvergence property:

REMARK 5.1 (SEMICONVERGENCE, [6]) *The total restoration error* $\rho_J = \|f_J -$

$f^{(0)}\|^2$ has a minimum for a certain value of J. Consequently, when J increases, the TSVD solution first approaches the true object $f^{(0)}$ and then goes away.

Proof. The *energy functional* for the class of TSVD solutions is given by:

$$E^2(f_J) = \|f_J\|^2 = \sum_{j=1}^{J} \frac{1}{\sigma_j^2} | < g, u_j > |^2 \tag{5.48}$$

and is an increasing function of J, hence a decreasing function of μ. The *discrepancy functional* is given by:

$$\epsilon^2(f_J, g) = \|Af_J - g\|^2 = \sum_{j=J+1}^{p} | < g, u_j > |^2 + \|u\|^2 \tag{5.49}$$

which means that it is decreasing with J and increasing with μ. The difference between a TSVD solution and the true object is:

$$f_J - f^{(0)} = \Psi_J g - f^{(0)} = (\Psi_J A f^{(0)} - f^{(0)}) + \Psi_J \nu \,. \tag{5.50}$$

It follows that the *approximation error* is:

$$\|\Psi_J A f^{(0)} - f^{(0)}\|^2 = \sum_{j=J+1}^{p} | < f^{(0)}, v_j > |^2 + \|v^{(0)}\|^2 \,, \tag{5.51}$$

where $v^{(0)}$ is the 'invisible' component of $f^{(0)}$, and the *noise-propagation error* is:

$$\|\Psi_J \nu\|^2 = \sum_{j=1}^{J} \frac{1}{\sigma_j^2} | < \nu, u_j > |^2 \,. \tag{5.52}$$

The approximation error decreases with J and the noise-propagation error increases with J. This demonstrates the existence of an optimal value J_{opt} for which the total restoration error $\rho_J = \|f_J - f^{(0)}\|$ attains a minimum. ■

Finding the optimal value J_{opt} is not a trivial problem. Methods to obtain it are described in section 5.5.7.

Positivity constraint and TSVD

The TSVD solution often has negative parts even when the physics dictates, as in the case of intensity map reconstruction from integral data, that the solution should be non-negative. The major question is how to incorporate this a priori information in the solution to best advantage.

De Villiers et al. ([95]) proposed a method based on the idea of adding in amounts of 'invisible' objects to remove the negative portions of the TSVD solution. The 'invisible' objects are elements of the null-space $\mathcal{N}(A)$, i.e. linear combinations of the singular functions removed from the generalised solution f^+ when the TSVD solution has been obtained. The solution is chosen, from the wide range of positive solutions, to be the one of minimum 2-norm. If the TSVD solution turns out non-negative, than this solution will coincide with it.

In order to obtain a solution $\hat{f}$ which agrees with the TSVD solution on the first J singular vector coefficients , the condition imposed is that these coefficients coincide with the known values $d_j =< g, u_j > /\sigma_j$ which appear in equation (5.46). This constraint is expressed by ([95]):

$$(V_J^T \hat{f})_j = d_j\,, \quad j = 1, \ldots, J\,, \tag{5.53}$$

where V_J is the matrix formed by the first J singular vectors $v_1, \ldots, v_J$, which are columns of the singular matrix V. The purpose of the method is therefore to find the minimal norm positive solution $\hat{f}$ whose decomposition in the basis $\{v_1, \ldots, v_p\}$ has the form ([95]):

$$\hat{f} = \sum_{j=1}^{J} d_j v_j + \sum_{j>J} a_j v_j = f_J + \sum_{j>J} a_j v_j\,. \tag{5.54}$$

Hence $\hat{f}$ is the sum between the TSVD solution f_J and an 'invisible' object. The determination of this solution is a mathematical programming problem expressed by ([95]):

$$\min_{f \geq 0,\, V_J^T f = d} \| f \|^2\,. \tag{5.55}$$

An alternative equivalent formulation is to denote the unknown f by $f = Va$ and

to seek a vector $a \in R^p$ which is the solution of the quadratic programming problem ([95]):

$$\min_{a_j=d_j,\, j=\overline{1,J},\, 0\leq Va<\infty} \|Va\|^2 . \tag{5.56}$$

The problem is consistent if the feasible set is non-empty. If it exists, the solution is unique since the norm is a strictly convex function minimised over a convex set. However, the existence of the solution is a non-trivial problem and there are certain values of d for which the feasible set is empty ([95]). Namely, when there exists $j \in \overline{1, J}$ such that $(Vd)_j < 0$, the programming problem has no solution.

5.5.4 Landweber Regularization and Variants

Landweber method

In 1951, Landweber proposed ([48]) an iteration formula for obtaining successive approximations for the solution of the Fredholm integral equation ([96]) of the first kind and proved convergence under various conditions. The method has been further studied and improved by Eicke ([28]), Strand ([86]), Defrise and de Mol ([22]), Bialy ([9]), Piana et al.([79]), and Bertero et al. ([6]).

The method is based on the fact that the set of solutions of the unconstrained equation (5.17) coincides with the set of fixed points of the following non-linear operator G ([48]):

$$G(f) = f + \tau(A^*g - A^*A\, f) . \tag{5.57}$$

REMARK 5.2 *If the relaxation parameter τ satisfies the condition:*

$$0 < \tau < \frac{2}{\|A\|^2} , \tag{5.58}$$

then the operator G has the property of being non-expansive, i.e:

$$\|G(f_1) - G(f_2)\| \leq \|f_1 - f_2\| , \tag{5.59}$$

for any f_1, f_2.

Proof. This condition is derived from the fact that ([48]):

$$\begin{aligned} \|G(f_1) - G(f_2)\| &= \|(f_1 - f_2)(I - \tau\, A^*A)\| \\ &\leq \|f_1 - f_2\| \cdot \|I - \tau\, A^*A\|\,. \end{aligned} \tag{5.60}$$

Using the singular value decomposition of the linear operator A: $A = U\,\Sigma\,V^T$, and denoting by σ_{max} the maximum singular value, then:

$$\|I - \tau A^*A\| = \|U\,(I - \tau\Sigma^2)U^T\| = \sqrt{1 - \tau\sigma_{max}} \tag{5.61}$$

From Eqs. (5.60) and (5.61), it follows that G is non-expansive if and only if $|1 - \tau\sigma_{max}| \leq 1$, which is equivalent to Eq. (5.58) because $\|A\| = \sigma_{max}$. ■

From the general fixed point theorems for non-expansive operators ([77]), if G has at least a fixed point, then the sequence $\{f_n\}_{n\geq 1} = \{G(f_{n-1})\}$ converges weakly to a fixed point of G. Moreover, Bialy ([9]) proved the following stronger result using spectral methods:

THEOREM 5.4 (BIALY, [9]) *If $g \in \mathcal{R}(A) \oplus \mathcal{R}(A)^{\perp}$ and τ satisfies conditions (5.58), then $\{f_n\}_{n\geq 1}$ converges strongly for any initial value f_0 and*

$$\lim_{n\to\infty} f_n = f^+ + P_{\mathcal{N}(A)} f_0\,, \tag{5.62}$$

where $P_{\mathcal{N}(A)}$ is the orthogonal projection onto $\mathcal{N}(A)$. In particular, the limit is f^+ when $f_0 = 0$.

The Landweber method is the iterative derivation of approximations of the fixed points of the operator G:

Algorithm 5.1 (Landweber, [48]) *1. Consider an arbitrary initial guess f_0.*

2. Compute recursively approximations of a fixed point according to the rule:

$$f_{k+1} = G(f_k) = f_k + \tau(A^*g - A^*A\, f_k)\,. \tag{5.63}$$

with τ satisfying the conditions: $0 < \tau < 2/\|A\|^2$.

REMARK 5.3 (SEMICONVERGENCE, [6]) *If the data is noisy, g generally does not satisfy the condition of the theorem, therefore the convergence does not occur. However, a so-called* semiconvergence *property stands: when k increases, the restoration error* $\rho_k = \|f_k - f^{(0)}\|$ *first decreases towards a minimum, then increases.*

Proof. In order to approach the semiconvergence property, the singular value decomposition of the operator A is used:

$$A\,f = (U\Sigma V^T)f = \sum_{j=1}^{n} \sigma_j < f, v_j > u_j\,, \tag{5.64}$$

where σ_j are the singular values and u_j, v_j are the singular vectors of A - the columns of the matrices U and V respectively.

With these notations, from equation (5.63), an iterate f_k can be represented as follows:

$$f_k = \sum_{j=1}^{n} \left(1 - (1 - \tau\sigma_j^2)^k\right) \frac{< g, u_j >}{\sigma_j} v_j \tag{5.65}$$

It follows that, when $k \to \infty$, the expression converges to the generalised solution $f^+ = \sum_{j=1}^{n} \frac{<g,u_j>}{\sigma_j} v_j$. This limit is not physically meaningful in the noisy case because the problem is ill-conditioned, therefore the regularization properties need to be investigated. The relevant quantities are ([6]):

- the *energy functional*:

$$E^2(f_k) = \|f_k\|^2 = \sum_{j=1}^{n} \frac{1}{\sigma_j}^2 |1 - (1 - \tau\sigma_j^2)^k|^2 | < g, u_j > |^2\,, \tag{5.66}$$

  which is an increasing function of k.

- the *discrepancy functional*:

$$\epsilon^2(f_k, g) = \|Af_k - g\|^2 = \sum_{j=1}^{n} (1 - \tau\sigma_j^2)^{2k} | < g, u_j > |^2 + \|u\|^2\,, \tag{5.67}$$

  which is a decreasing function of k.

If $\Psi^{(k)}$ denotes the linear operator which maps g into f_k, defined by (5.63), such that $f_k = \Psi^{(k)} g$, then the difference between the reconstruction f_k and the true,

noiseless object $f^{(0)}$ is:

$$f_k - f^{(0)} = (\Psi^{(k)} A\, f^{(0)} - f^{(0)}) + \Psi^{(k)} \nu\,. \tag{5.68}$$

Hence, the *approximation error* is ([6]):

$$\|\Psi^{(k)} A\, f^{(0)} - f^{(0)}\|^2 = \sum_{j=1}^{n} (1 - \tau\sigma_j^2)^{2k} |< f^{(0)}, v_j >|^2 \tag{5.69}$$

and the *noise-propagation error* is ([6]):

$$\|\Psi^{(k)} \nu\|^2 = \sum_{j=1}^{n} \frac{1}{\sigma_j^2} |1 - (1 - \tau\sigma_j^2)^k|^2 |< g, u_j >|^2\,. \tag{5.70}$$

The approximation error is a decreasing function of k, whilst the noise-propagation error increases with k. Therefore, the restoration error $\rho_k = \|f_k - f^{(0)}\|$ has a minimum for a certain value k_{opt} and the semiconvergence property holds true. ■

Projected Landweber Method

The projected Landweber method has been derived from the Landweber method in order to tackle the constrained least-squares problem (5.20). Its regularization and convergence properties have been studied by Eicke ([28]). The method can be easily implemented if the projection onto the convex set $\mathcal{C}$, $P_{\mathcal{C}}$, is easily computable. Many numerical simulations and practical applications have shown that this method can provide much better estimates than the usual linear regularization methods. The practical drawback is that the convergence is too slow, that is a large number of iterations are required.

In order to summarise the convergence properties of the method, the following notations are considered:

- $I = \overline{A(\mathcal{C})}$ is the closure of the set $A(\mathcal{C})$, which is also convex.
- $P_{\mathcal{I}}$ is the convex projection onto $\mathcal{I}$.

THEOREM 5.5 *The constrained problem (5.20) has a solution if and only if* $P_{\mathcal{I}} g \in A(\mathcal{C})$. *In this case, the set of constrained least-squares solutions is closed and convex ([28]).*

As a consequence of this result, it follows that there exists a constrained least-squares solution of minimal norm, denoted by f_C^+. Eicke proved ([28]) that the set of the constrained least-squares solutions coincides with the set of the fixed points of the nonlinear operator $P_C G$, given by:

$$P_C G(f) = P_C(f + \tau(A^* g - A^* A\, f))\,. \tag{5.71}$$

In general, the projection operator P_C is nonlinear and non-expansive. It follows that, for τ satisfying the conditions 5.58, the compound operator $P_C G$ is also non-expansive and the projected Landweber method is the iterative derivation of approximations to the fixed points of the operator $P_C G$:

Algorithm 5.2 (Projected Landweber method, [28], [6])

1. *Consider an arbitrary initial guess* f_0.

2. *Compute recursively approximations of a fixed point according to the rule:*

$$\overline{f}_{n+1} = P_C G(\overline{f}_n) = P_C(\overline{f}_n + \tau(A^* g - A^* A\, \overline{f}_n))\,, \tag{5.72}$$

with $0 < \tau < 2/\|A\|^2$.

From the properties of the operator G and general fixed point theorems for non-expansive operators, the following **convergence** result has been derived ([28]):

THEOREM 5.6 *If* $P_{\mathcal{I}} g \in A(C)$, *then the sequence* $\{\overline{f_n}\}_{n\geq 1}$ *given by 5.72 converges weakly to a fixed point of* $P_C G$, *which is a solution of the constrained problem (5.20).*

There are many open problems which concern the convergence of the projected Landweber method. As emphasised in [79], although an analytical proof has not yet been derived, numerical simulations strongly indicate that the **semiconvergence property** holds in the case of noisy data.

Preconditioned Landweber and projected Landweber method

The Landweber and projected Landweber methods have the disadvantage of requiring a large number of iterations until convergence is attained. Therefore, efforts

have been made ([86], [82], [79]) to modify the iteration structures such that, simultaneously, the regularization properties be kept and the convergence rate be improved.

A preconditioned (or generalised) variant of the Landweber iteration has been reported by Strand ([86]). It is based on the use of a linear operator $H : \mathcal{R}(A^*) \rightarrow \mathcal{R}(A^*)$, which is continuous, positive definite and bounded from below (hence, has a continuous inverse), and commutes with A^*A. Using the operator H, whose inverse H^{-1} does exist, the Euler equation 5.18 has the same set of solutions as the following *preconditioned* equation:

$$HA^*Af = HA^*g\,, \tag{5.73}$$

and the corresponding iteration is:

$$\hat{f}_{k+1} = \hat{f}_k + \tau H(A^*g - A^*A\hat{f}_k)\,, \tag{5.74}$$

with τ satisfying the conditions:

$$0 < \tau < \frac{2}{\|HA^*A\|}\,. \tag{5.75}$$

Strand's analysis concentrates on a special class of preconditioners H defined by the following behaviour with respect to the singular vectors v_l (the columns of the matrix V of the singular value decomposition $A = U\Sigma V^T$):

$$Hv_l = p_l v_l\,, \tag{5.76}$$

where $0 < p_l\lambda_l < 2$ for all the singular values λ_l of the operator A^*A.

The main **convergence** theorem for the preconditioned Landweber method using this class of preconditioners may be summarised by the statement that $\hat{f}_k$ converges strongly as $k \rightarrow \infty$ if and only if the projection of g onto $\mathcal{U} = sp\{u_1, \ldots, u_n\}$ is in the range of A, $\mathcal{R}(A)$. If $\hat{f}_k$ does not converge, then $\|\hat{f}_k\| \rightarrow \infty$ as $k \rightarrow \infty$. This is precisely the Bialy Theorem 5.62 for the preconditioned Landweber method ([86], [6]). Moreover, it can be proved that, if (5.74) converges to $\tilde{f}$ for one starting function in $\mathcal{V} = sp\{v_1, \ldots, v_n\}$, it converges to $\tilde{f}$ for any starting function in $\mathcal{V}$.

Also, the sequence $\hat{f}_k$ has the **semiconvergence** property in the case of noisy data ([79]).

In practice, an extremely simple method of achieving (5.76) is to take $H = F(A^*A)$, where $F(.)$ is a polynomial or rational function such that the function $\phi(\lambda) = \lambda F(\lambda)$ mimics the behaviour of the Heaviside step function ([79], [83]). Examples of such operators are provided in Strand's paper [86].

Another example of a preconditioner is given by Sanz and Huang ([82]) as $H = (A^*A + \gamma I)^{-1}$, with $\gamma > 0$, which assures that $A^*A + \gamma I$ is always invertible. In this case, the function $\phi(\lambda)$ is:

$$\phi(\lambda) = \frac{\lambda}{\lambda + \gamma} \tag{5.77}$$

and the condition 5.75 on the relaxation parameter becomes:

$$0 < \tau < \frac{2(\|A\|^2 + \gamma)}{\|A\|^2}\,. \tag{5.78}$$

The accuracy and convergence rate of the reconstruction are sensitive to the choice of H. As shown in [86], [82], and supported by the numerical simulations presented in section 6, the method requires less iterations than the classical Landweber method.

The Projected Landweber method can be accelerated in a similar way ([79]) if the following iterative scheme:

$$\bar{f}_{k+1} = P_C(\bar{f}_k + \tau H(A^*g - A^*A\bar{f}_k))\,. \tag{5.79}$$

is considered.

Piana and Bertero ([79]) provide another example of an operator H. They begin with the observation that a method of accelerating the Projected Landweber method is to apply the projection onto $\mathcal{C}$ not at each iteration, but every $N \geq 2$ iterations and show that this procedure can be described in terms of a preconditioner H which is a polynomial of degree $N-1$. Their results indicate that this latter preconditioner allows considerable acceleration of the Projected Landweber method at the cost of a moderate increase of the minimum restoration error, and that, in general, the acceleration depends on the amount of noise affecting the data. However, in the

examples studied in Chapter 6, Sanz's preconditioner performs more efficiently than Piana's preconditioner.

5.5.5 Steepest Descent Method

The steepest descent method is a regularization method derived from the gradient descent minimisation of the discrepancy functional $\epsilon^2(f,g) = \|Af - g\|^2$ ([6]). The discrepancy functional can be expanded as:

$$\epsilon^2(f,g) = \|Af-g\|^2 =< A^*Af, f > -2 < A^*g, f > +\|g\|^2 = 2\eta(f,g)+\|g\|^2 \quad (5.80)$$

where $\eta(f,g)$ denotes the functional:

$$\eta(f,g) = (1/2) < A^*Af, f > - < A^*g, f > \ . \quad (5.81)$$

g is the given image, therefore $\|g\|^2$ is a constant, and minimising the discrepancy $\epsilon(f,g)$ is equivalent to minimising the functional $\eta(f,g)$. The gradient descent theorem ([77],[5]) asserts that, given an approximation f_k of the object, in the neighbourhood of this point, the functional $\eta(f,g)$ decreases most rapidly in the direction of the negative gradient $-\nabla_f \eta(f,g)$. It can be easily seen from (5.80) that this gradient is equal to half of the gradient of the discrepancy, which means:

$$\nabla_f \eta(f,g) = \frac{1}{2}\nabla_f \epsilon(f,g) = A^*(Af - g)\,. \quad (5.82)$$

Hence, a better approximation of the real object (and of the minimum discrepancy) is obtained via the following iteration ([6]):

$$f_{k+1} = f_k - \tau \nabla_f \eta(f,g)\,, \quad (5.83)$$

with $\tau > 0$. In this case, the value of $\eta(f_{k+1}, g)$ is a quadratic polynomial function of τ:

$$\eta(f_{k+1},g) = \eta(f_k,g) + \frac{1}{2}\tau^2 \|AA^*(Af_k - g)\|^2 - \tau\|A^*(Af_k - g)\|^2\,. \quad (5.84)$$

The optimal value for the parameter τ is obviously the one which gives a minimal $\eta(f_{k+1}, g)$ and is obtained from the previous equation as ([6]):

$$\tau_k = \frac{\|A^*(Af_k - g)\|^2}{\|AA^*(Af_k - g)\|^2}\,. \tag{5.85}$$

Hence, the following steepest decent algorithm can be expressed.

Algorithm 5.3 (Steepest descent) *1. Consider an arbitrary initial guess f_0.*

2. Compute recursively approximations of the optimal solution according to the rule:

$$f_{k+1} = f_k - \tau_k A^*(Af_k - g)\,. \tag{5.86}$$

with τ_k given by:

$$\tau_k = \frac{\|A^*(Af_k - g)\|^2}{\|AA^*(Af_k - g)\|^2}\,. \tag{5.87}$$

The interpretation of equation (5.86) is that the steepest descent method is a variant of the Landweber method with the steplength of the relaxation parameter τ optimised.

REMARK 5.4 (CONVERGENCE AND SEMICONVERGENCE) *It has been proven ([36], [41], [42], [43]) that, if $f_0 = 0$, the iterates f_k* converge *to f^+ in the case of a noise-free image and that they have the* semiconvergence property *in the case of noisy images, i.e. there exists an optimum number of iterations which provides the best approximation of the unknown object.*

Stopping rules for the choice of the optimal number of iterations in the noisy case are presented in section 5.5.7.

5.5.6 Conjugate Gradient Method

The conjugate gradient method, reported as a general numerical analysis method ([6],[77],[5]), is expressed by the following iterative scheme:

Algorithm 5.4 (Conjugate Gradient method) *1. Consider the initial values $f_0 = 0$, $r_0 = p_0 = A^*g$.*

2. Compute the iterated values:

$$\begin{aligned} \alpha_k &= \frac{\|r_k\|^2}{\|Ap_k\|^2} & r_{k+1} &= r_k - \alpha_k A^* A p_k \\ \beta_k &= \frac{\|r_{k+1}\|^2}{\|r_k\|^2} & p_{k+1} &= r_{k+1} + \beta_k p_k \\ f_{k+1} &= f_k + \alpha_k p_k \end{aligned} \tag{5.88}$$

until convergence.

REMARK 5.5 (CONVERGENCE AND SEMICONVERGENCE, [6]) *In the noise-free discrete case, convergence is attained in a number of steps equal to the dimension n of the data space if the vectors* $A^*g, \ldots, (A^*A)^{n-1}A^*g$ *are linearly independent. In the infinite-dimensional case,* f_k *converges to* f^+ *when* $k \to \infty$ *if* $g \in \mathcal{R}(A)$. *The method has the semiconvergence property in the case of a noisy image g.*

The method can be applied assuming that the vectors $A^*g, \ldots, (A^*A)^{k-1}A^*g$ are linearly independent. The iterate solution f_k represents a minimum point of the discrepancy functional ϵ in the k-th Krylov space $\mathcal{K}^{(k)} = sp\{A^*g, \ldots, (A^*A)^{k-1}A^*g\}$. The convergence is faster than that obtained with the steepest descent method ([43]).

As stressed throughout the chapter, a proper use of the semiconvergence property requires rules to estimate the value of the regularization parameter which provides a minimum of the restoration error. Such rules are summarised in section 5.5.7.

5.5.7 Stopping rules and parameter choice for regularization methods

For ill-posed problems, a compromise has to be found between accuracy and stability, and the regularization algorithms must be completed by means of a rule for choosing the optimal regularization parameter, which can be either $\mu_{opt} > 0$ for Tikhonov method, or the optimum index J_{opt} for the truncation of the SVD, or the optimal number of iterations k_{opt} for the iterative methods. In the following, μ will denote a generic regularization parameter, which can be any of the ones already enumerated.

The optimal regularization parameter μ_{opt} is the one for which the regularized solution $f_{\mu_{opt}}$ has minimal distance to the true object $f^{(0)}$ ([6]). In practice, the true object is not known, so the optimal value needs to be inferred from other properties,

like the level of noise in the image or the energy of the object, or to be generated by the problem data themselves ([6],[67],[22], [47]). The following methods for achieving this are discussed: prescribing the energy of the object space ([6]), prescribing the discrepancy (the discrepancy principle [67]), Miller's method ([65], [6]), and cross-validation ([96]).

1. *Regularized solutions with prescribed energy ([6])*

   If the energy of the true object $E = \|f^{(0)}\|$ is known, then an estimate μ_1 of the optimal regularization parameter is the one for which the solution has the same energy as the true object:

$$\|f_{\mu_1}\| = E\,. \tag{5.89}$$

2. *Regularized solutions with prescribed discrepancy - The discrepancy principle [67], [92]*

   If the energy of the noise in the image, $\epsilon = \|\nu\| = \|Af^{(0)} - g\|$, is known, than an estimate μ_2 of the optimal parameter is the one which provides a solution which has the discrepancy equal to ϵ:

$$\|Af_{\mu_2}\| = \epsilon\,. \tag{5.90}$$

In the case of the Landweber and TSVD methods, the discrepancy principle has been further refined by Defrise and de Mol in [22]. They showed that, in the case of the Landweber method, an estimate of the optimal regularization parameter is obtained for an iteration number K which satisfies:

$$\begin{aligned} \|Af_K - g\|^2 &> b\epsilon^2 \\ \|Af_{K+1} - g\|^2 &\leq b\epsilon^2 \end{aligned} \tag{5.91}$$

with $b = 2/(2 - \tau\|A\|^2)$. For the TSVD method, they propose the use of a truncation of the SVD which contains only the singular values whose corre-

sponding singular vectors u_j satisfy the condition:

$$| < g, u_j > | \geq 2\epsilon \,. \tag{5.92}$$

They assert that these latter refinements provide good estimates of the optimal regularization parameters.

3. *Miller's method ([65], [6])*

An approach to ill-posed problem proposed by Miller ([65]) can also be used as a method for estimating the optimal regularization parameter ([6]). It requires that a bound E of the energy and a bound ϵ of the discrepancy of the unknown object are known. If this is so, then all the admissible approximate solutions f satisfy:

$$\|Af - g\| \leq \epsilon \,, \quad \|f\| \leq E \,. \tag{5.93}$$

If the above set of equations has at least a solution, then it has been shown ([65]) that the value:

$$\mu_3 = \left(\frac{\epsilon}{E}\right)^2 \tag{5.94}$$

is an estimate of the optimal parameter. However, numerical simulations have indicated ([6]) that Miller's method tends to underestimate the optimal value of the regularization parameter for Tikhonov regularization and to overestimate the optimal number of iterations for the iterative methods when the exact values of ϵ and E are used.

4. *Cross-validation ([96], [18])*

Another method, called *weighted cross-validation*, for choosing the optimal parameter has been reported by Craven and Wahba ([18], [96]) in the case of the Tikhonov regularization. It can be also applied to iterative methods (Landweber, steepest descent, conjugate gradient), inasmuch as the inverse of the number of iterations $\mu = 1/k$ acts as the regularization parameter of the Tikhonov method. The objective of the cross-validation is to derive the estimate of the optimal parameter from the data, without additional knowledge or prediction of the energy or noise level. The fundamental idea is that, if $f_{\mu,k}$

is the minimiser of the problem where the kth data component is missing, that is:

$$\frac{1}{n}\sum_{j=1\,,j\neq k}^{n}\left((Af)_j - g_j\right)^2 + \mu\|f\|^2\,, \tag{5.95}$$

and if $\hat{\mu}$ is a good choice for the regularization parameter, then $Af_{\hat{\mu},k}$ should be closer to g_k, on average, than $Af_{\mu,k}$ for other values of μ (i.e. a good value of the regularization parameter would predict any missing data values). The measure of this closeness is the following weighted mean square data prediction error ([96]):

$$V_0(\mu) = \frac{1}{n}\sum_{k=1}^{n}\left((Af_{\mu,k})_k - g_k\right)^2 w_k(\mu)\,, \tag{5.96}$$

where $w_k(\mu)$ are some weights given in [96]. The choice $\hat{\mu}$ is the minimiser of $V(\mu)$.

It has been further proven ([18],[6]) that $V_0(\mu)$ can be computed using the simpler expression:

$$V_0(\mu) = \frac{1}{n}\sum_{k=1}^{n}\frac{\left|(Af_\mu)_k - g_k\right|^2}{\left|1 - A_{kk}(\mu)\right|^2}\,, \tag{5.97}$$

which does not require the solutions of n minimisation problems. In the equation, f_μ is the regularized solution and $A(\mu) = AA^*(AA^* + \mu I)^{-1}$.

The cross-validation function V_0 has the drawback of not being invariant with respect to linear transformation of the components of the image or object, therefore, it is replaced by the *generalised cross-validation (GCV) function* ([18]):

$$V(\mu) = \frac{\|Af_\mu - g\|^2}{(Tr(I - A(\mu)))^2}\,, \tag{5.98}$$

where:

$$Tr(I - A(\mu)) = \sum_{j=1}^{n}[1 - A_{jj}(\mu)\,. \tag{5.99}$$

Figure 5.8 presents the GCV function $V(\mu)$ at the Tikhonov regularization of the reconstruction of a nine point object from a noisy integral image (Chapter 6, Example 6.2).

The GCV function has another important property ([18],[6]), namely, its minimiser μ_{GCV} provides an estimate of the minimiser μ_4 of the *predictive mean*

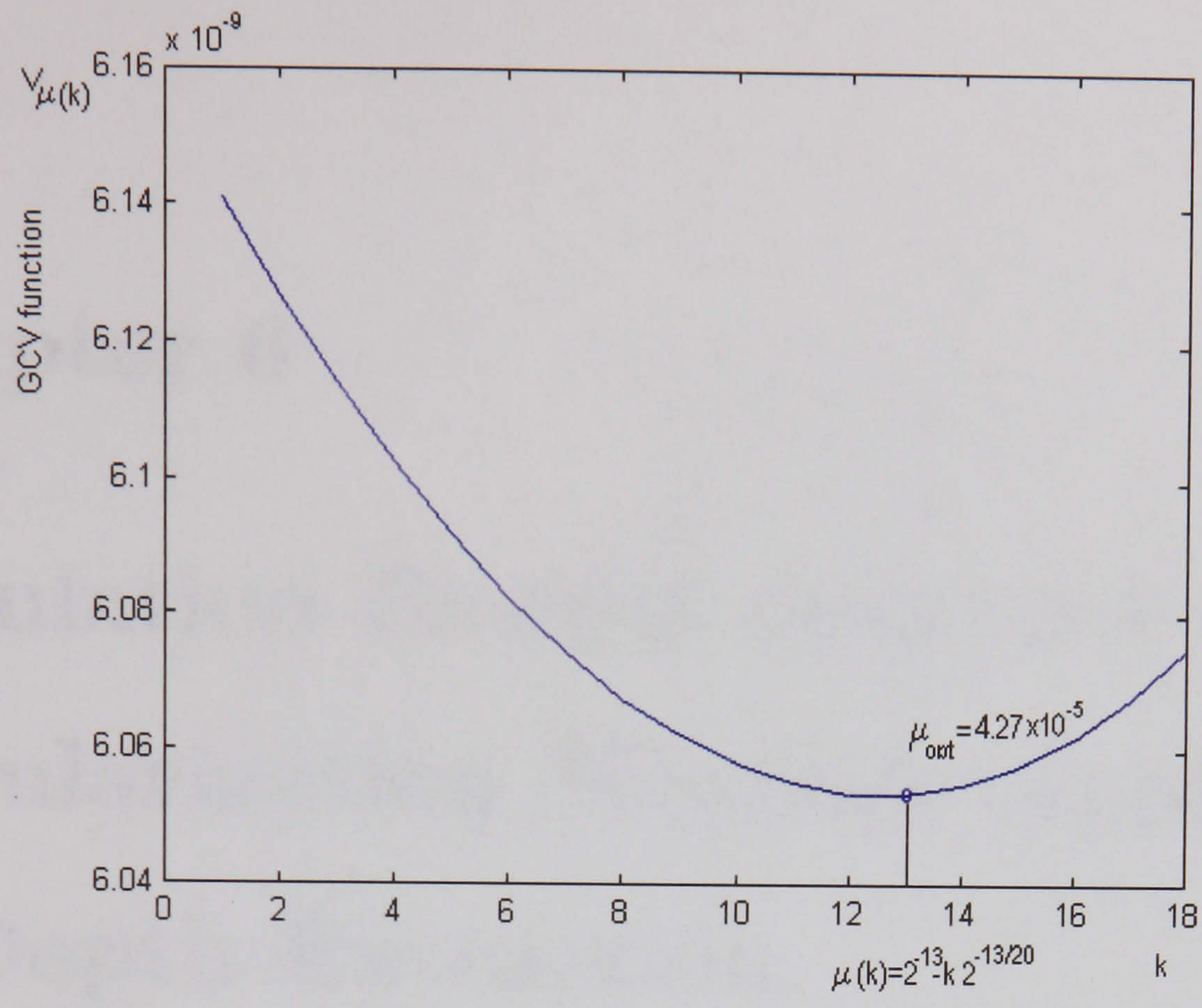


Figure 5.8: The Generalised Cross Validation function at the Tikhonov regularization of the reconstruction of a nine point object from an integral image (Chapter 6, Example 6.2).

square error:

$$T(\mu) = \|Af_\mu - g^{(0)}\|, \tag{5.100}$$

where $g^{(0)}$ is the noise-free image. It is worth mentioning that the minimiser μ_4 of $T(\mu)$, which is unique, provides another estimate of the optimal regularization parameter ([6]).

5.6 Conclusions

Depth extraction has been formulated as an inverse problem of the image formation process, which has proven very ill-conditioned. To overcome the ill-conditioning, regularization methods are used to obtain physically meaningful solutions of the problem. The section has presented an extensive overview of existing regularization methods and their convergence properties. Also, methods of choosing the optimal regularization parameters in conditions of noise have been presented.

In the following chapter, these methods are used in simulations to provide realistic solutions to the ill-posed inverse problem of object space reconstruction, and, subsequently, depth extraction from integral images.

Chapter 6

Simulation Results obtained with Regularization Methods Applied to Depth Extraction

6.1 Introduction

In Chapter 5, depth extraction has been formulated as an inverse problem and regularization methods able to provide realistic solutions have been reviewed. In the following, these methods are applied to a set of unidirectional computer generated integral images, and the reconstructions of the object space thus obtained are assessed.

The following procedure has been followed:

1. For each computer generated image, noisy representatives were generated using white Gaussian noise.

2. The imaging matrix A corresponding to a sampling grid covering a relevant portion of the object space was computed using the mathematical model presented in Chapter 3 and discretised version presented in Chapter 5.

3. The object space was then reconstructed using appropriate regularization methods, and, for each method, the optimal solution has been computed. The quality of the reconstruction has been assessed using the relative restoration

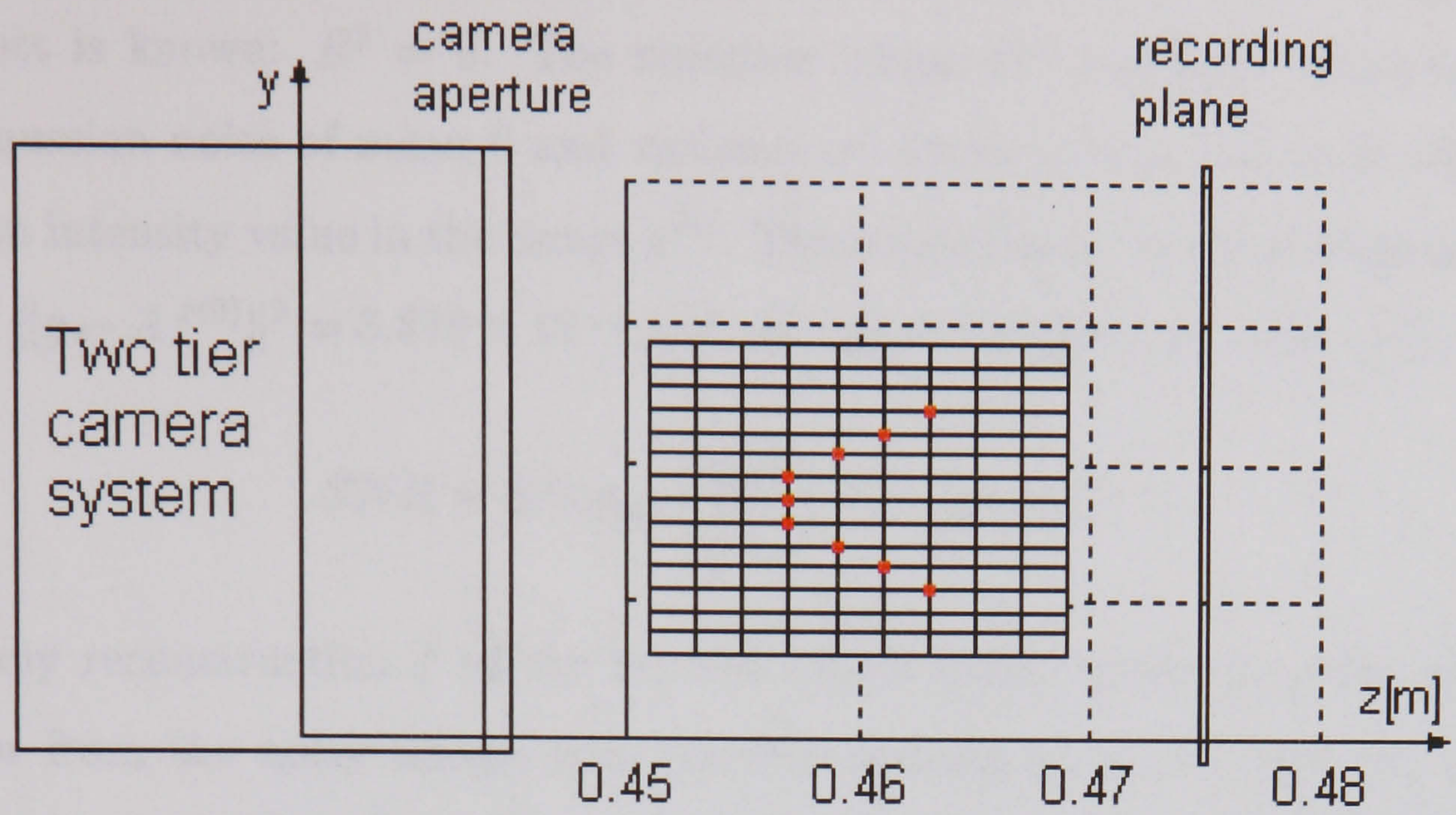


Figure 6.1: The object space configuration corresponding to the integral image of nine luminous points considered in Example 6.2.

error $\rho_{rel} = \|f_\mu - f^{(0)}\|/\|f^{(0)}\|$ in the ideal case and using an error estimator in non-ideal cases.

4. For each regularization method, various optimal regularization parameter estimations have been considered and the results have been compared.

6.2 Example 1. Reconstruction of a nine point object

The first example is a 'toy' situation shown in Figure 6.1. The aim is to recover the object space from a computer generated picture $g^{(0)}$ representing a set of nine luminous points situated at various (known) depths and lateral positions in space. The integral image is shown in Figure 8.2, Appendix C. All nine object points are situated in the same horizontal plane, as shown in Figure 6.1.

For this purpose and in order to calculate an accurate relative restoration error, the object space has been sampled with a 150 point grid which contains the exact positions of the nine object points. Hence, the noiseless object $f^{(0)}$ is known with full precision: it is a function which is equal to 1 at the grid points coinciding with the luminous points and 0 at any of the other 141 points of the grid. The matrix A corresponding to the integral imaging operator for this sampling grid

has been computed as described in Chapter 5. For this image, the energy of the real object is known: $E^2 = 9$. The noiseless image $g^{(0)}$ has been corrupted with white Gaussian noise of mean 0 and variance σ^2 whose square root σ is 1% of the maximum intensity value in the image $g^{(0)}$. The level of noise in the corrupted image g is $\epsilon^2 = \|g - Af^{(0)}\|^2 = 3.878 \times 10^{-4}$ and the resulting signal to noise ratio is:

$$SNR = 20 \log_{10} \left(\frac{\|g\|}{\|\nu\|} \right) = 31.47\,(dB) \tag{6.1}$$

For any reconstruction f of the 9-point object space either from the noiseless image, or from the noisy image, it is possible accurately to compute the relative restoration error $\rho_{rel} = \|f - f^{(0)}\|/\|f^{(0)}\|$ (Figure 5.7), the energy of the resultant object $E^2(f) = \|f\|^2$ and to estimate the discrepancy $\epsilon^2(f, g) = \|Af - g\|^2$ (Figure 5.5).

Given all this, the enumerated regularization methods have been tested to produce object space reconstructions both from the noise-free image $g^{(0)}$ and from the noisy image g.

6.2.1 Noiseless image

The object space reconstruction results obtained from the noise-free image $g^{(0)}$ for various methods are shown in Table 6.1. Figure 6.2 presents the comparison between regularized reconstructions and the original intensity of the real object. Both the reconstructions (continuous red line) and the real object intensity (dashed green line) are plotted as vectors.

The most accurate results, exhibiting practically no restoration error, were provided by Tikhonov regularization. Choosing the regularization parameter is simple in the case of a noiseless image, but becomes a much more delicate and computationally intensive task in the noisy case. The solution provided by TSVD has the largest relative error, compared to the other methods. Moreover, the quadratic programming method which allows the derivation of a positive solution from the TSVD solution has not proven feasible for the test image. Of the iterative methods, the fastest, computationally least expensive are the Preconditioned Landweber and Projected Landweber methods with a Sanz preconditioner and the Conjugate Gradient

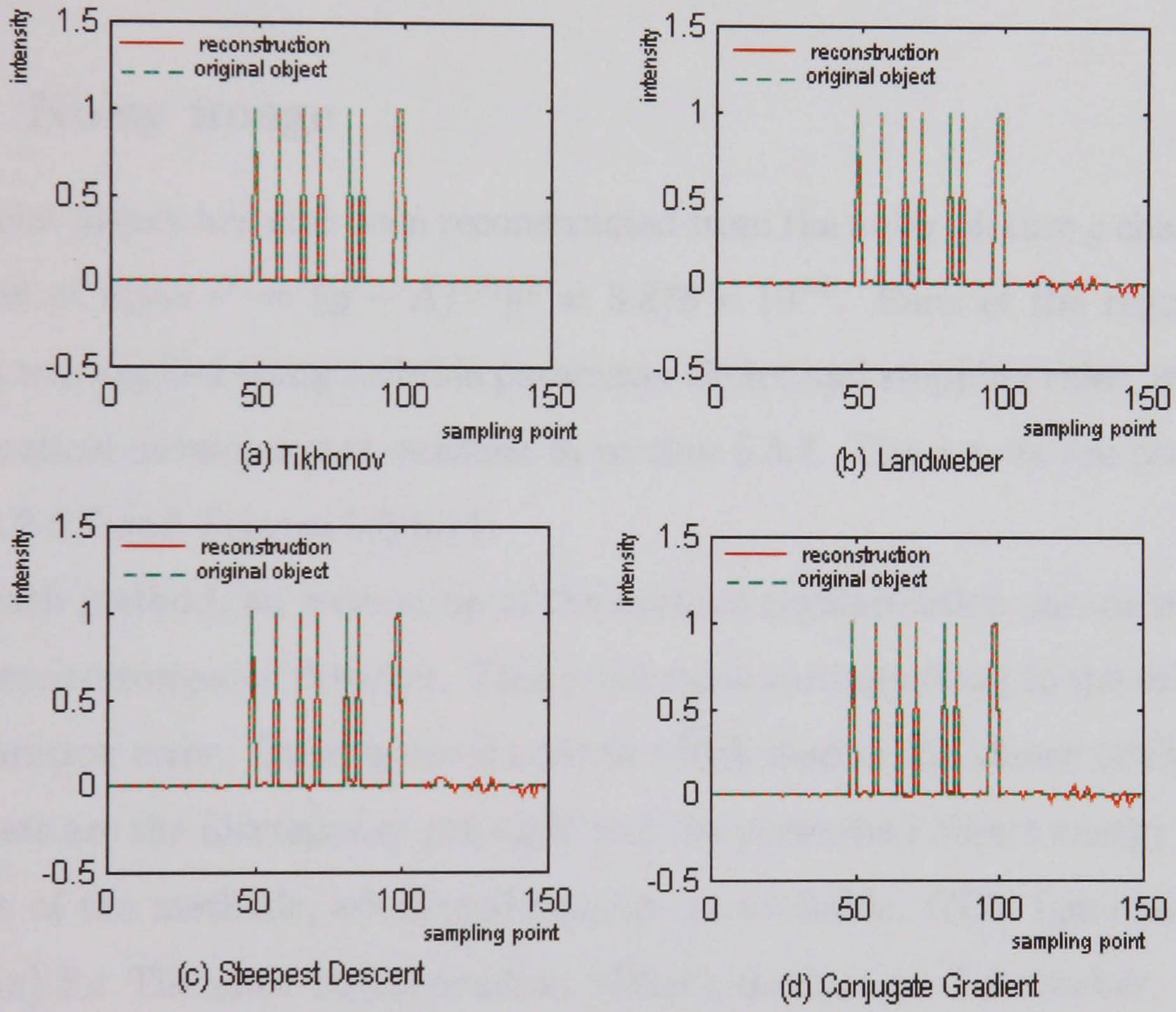


Figure 6.2: Reconstructions (continuous red line) of a *nine point object*, obtained with various regularization methods, from a *noiseless* integral image versus the real object intensity (dashed green line).

method. They also gave good accuracy with the relative restoration error ranging between 2.58% and 5.87%. In addition, the Projected Landweber method and the Preconditioned Landweber/Projected Landweber with a Piana preconditioner provided very good accuracy as well, but were relatively slow due to the large number of iterations involved. Compared to Landweber, Projected Landweber and Preconditioned Projected Landweber (Piana), the Steepest Descent method requires a smaller number of iterations, but is more computationally intensive and slower. This is due to the need of computing the optimal steplength of the relaxation parameter at each iteration.

6.2.2 Noisy image

The 9-point object has also been reconstructed from the noisy picture g characterised by a level of noise $\epsilon^2 = \|g - Af^{(0)}\|^2 = 3.878 \times 10^{-4}$. Each of the regularization methods was applied using suitable parameter choice and stopping rules, in line with the theoretical developments outlined in section 5.5.7. The results are presented in Tables 6.2-6.7 and Figures 6.3-6.11.

For each method, an estimation of the optimal regularization parameter is given via the semiconvergence theorem. This is the value corresponding to the minimum of the restoration error. Other general criteria which lead to the choice of the optimal parameters are the discrepancy principle and the prescribed object energy principle. For some of the methods, additional criteria are available: GCV (generalised cross validation) for Tikhonov regularization, Miller's method for Landweber, Tikhonov and Steepest Descent, Defrise's extension of the discrepancy principle for Landweber and TSVD.

Tikhonov regularization provided the results shown in Table 6.2 and Figure 6.3. All the parameter choice rules provided homogeneous and consistent results, the relative restoration error ranging from 4.99% to 7%. It is worth mentioning that the generalised cross validation method, which is the only criterion which provides the optimal parameter without any prior knowledge or assumption on the noise level in the image or on the energy of the object space, has the drawback of being computationally expensive and extremely slow. The solutions obtained with Tikhonov regularization, shown in Figure 6.3 versus the real object, have the disadvantage of

Method	Regularization parameter	Relative restoration error (ρ_{rel})	Estimated discrepancy ($\epsilon^2(f, g^{(0)})$)	Estimated object energy (E^2)
Tikhonov	$\mu = 4.76 \times 10^{-7}$	$7 \times 10^{-5}\%$	1.39×10^{-10}	8.99
TSVD	$J = 123$	8.37%	4.47×10^{-5}	8.93
TSVD with positivity	infeasible	–	–	–
Landweber	$k = 1656$	8.1%	9.4×10^{-5}	8.71
Projected Landweber	$k = 1042$	2.79%	4.98×10^{-5}	8.80
Preconditioned Landweber (Sanz, $\gamma = 0.1$)	$k = 138$	6.4%	4.94×10^{-5}	8.79
Preconditioned Projected Landweber (Sanz, $\gamma = 0.1$)	$k = 85$	2.58%	4.9×10^{-5}	8.82
Preconditioned Projected Landweber (Piana, $N = 5$)	$k = 1064$	2.9%	4.97×10^{-5}	8.79
Preconditioned Projected Landweber (Piana, $N = 10$)	$k = 1088$	2.98%	4.96×10^{-5}	8.78
Steepest descent	$k = 530$	6.51%	4.98×10^{-5}	8.79
Conjugate gradient	$k = 36$	5.87%	4.76×10^{-5}	8.90

Table 6.1: Regularization results for object space intensity reconstruction from a *noiseless* integral image containing *nine separate luminous points*. For each method, the regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

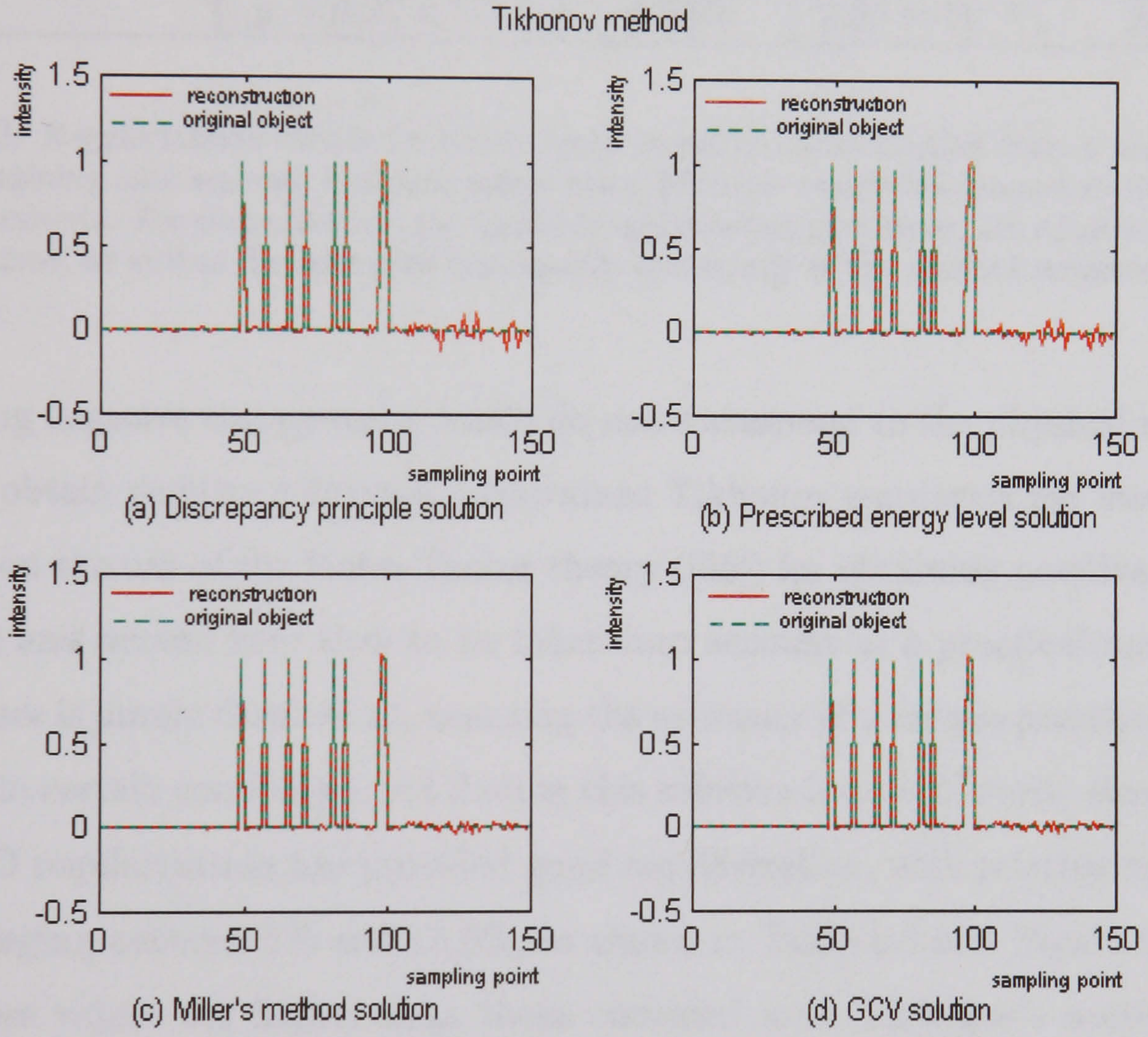


Figure 6.3: Reconstructions (continuous red line) of a *nine point object*, obtained with Tikhonov regularization method and associated parameter choice criteria, from a *noisy* integral image, versus the real object intensity (dashed green line).

Tikhonov regularization	Optimal regularization parameter	Relative restoration error (ρ_{rel})	Estimated discrepancy ($\epsilon^2(f, g^{(0)})$)	Estimated object energy (E^2)
Semiconvergence theorem	$\mu = 6.103 \times 10^{-5}$	4.99%	2.05×10^{-4}	8.92
Discrepancy principle	$\mu = 4.88 \times 10^{-4}$	7%	3.17×10^{-4}	8.48
Prescribed energy method	$\mu = 6.103 \times 10^{-5}$	4.99%	2.06×10^{-4}	8.901
Miller's method	$\mu = 4.35 \times 10^{-5}$	4.99%	2.05×10^{-4}	8.92
GCV	$\mu = 4.27 \times 10^{-5}$	4.99%	2.05×10^{-4}	8.92

Table 6.2: Regularization results for object space intensity reconstruction from a *noisy* integral image containing *nine separate luminous points* using Tikhonov regularization and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

containing negative components, which do not correspond to the physical reality. In order to obtain positive solutions, constrained Tikhonov regularization was tried. It is based on the use of the Kuhn-Tucker theory ([55]) for obtaining positive bounded solutions and proved very slow to be taken into account as a practical method. Its importance is purely theoretical, asserting the existence of a unique positive bounded solution in certain conditions, but finding this solution is an arbitrarily slow process.

TSVD regularization has provided good regularization, with relative restoration errors ranging between 5% and 10.6%, as shown in Table 6.3 and Figure 6.4. However, these values are higher than those obtained with Tikhonov's method. The explanation can be found in the fact that the Tikhonov regularization parameter μ is chosen from a continuous domain, whilst TSVD is confined to a finite, even low, number of truncation options. Therefore, the regularization gain cannot attain the precision of the Tikhonov method.

In an attempt to obtain a positive solution of the problem, TSVD with positivity constraint was tried. It has proved useless for the given image, there being no feasible solutions. The reason is that the constraints are overly stringent as there are negative values $(V\,d)_j$, therefore the initial condition required by the quadratic programming problem (5.56) is not fulfilled.

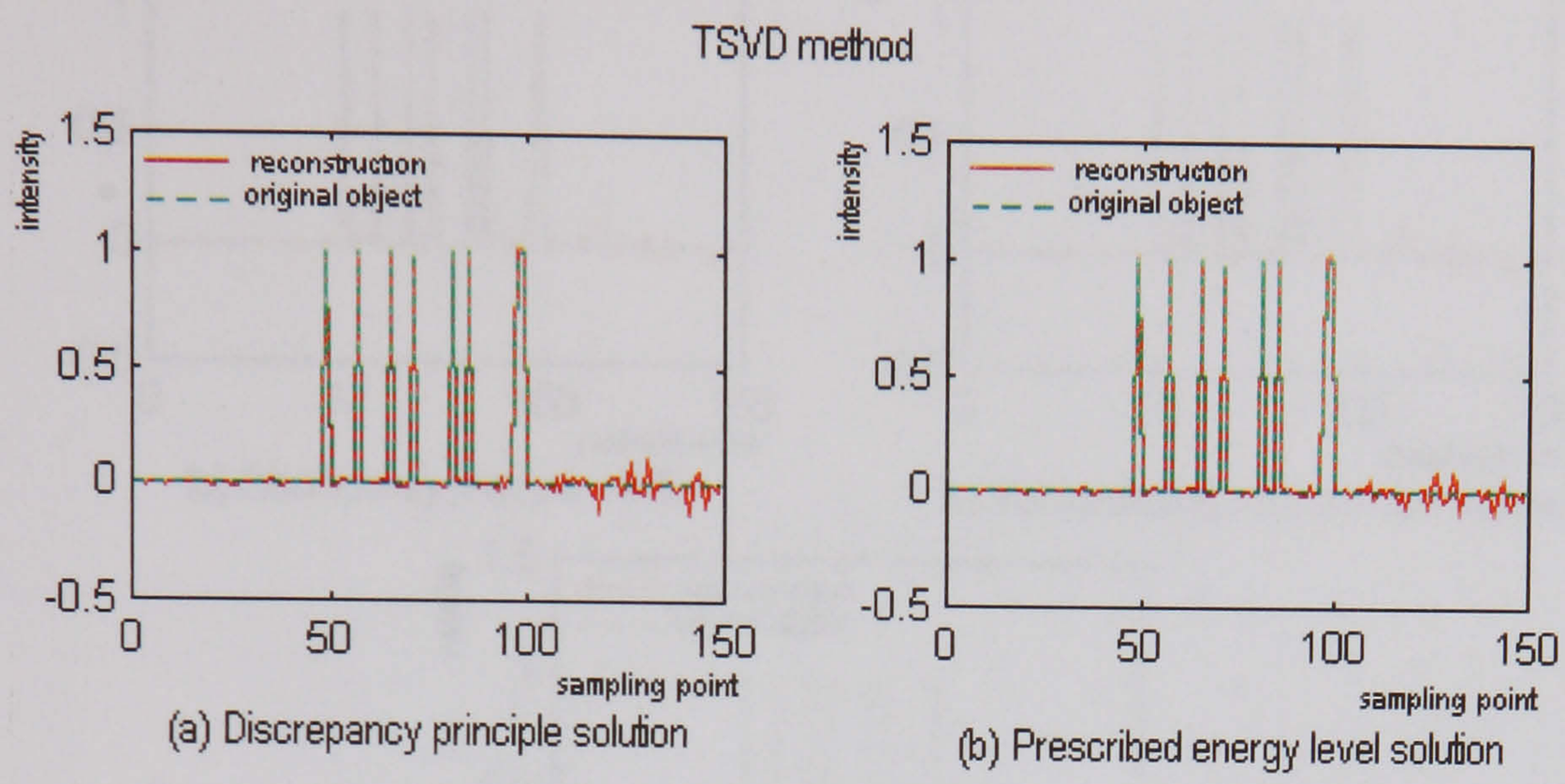
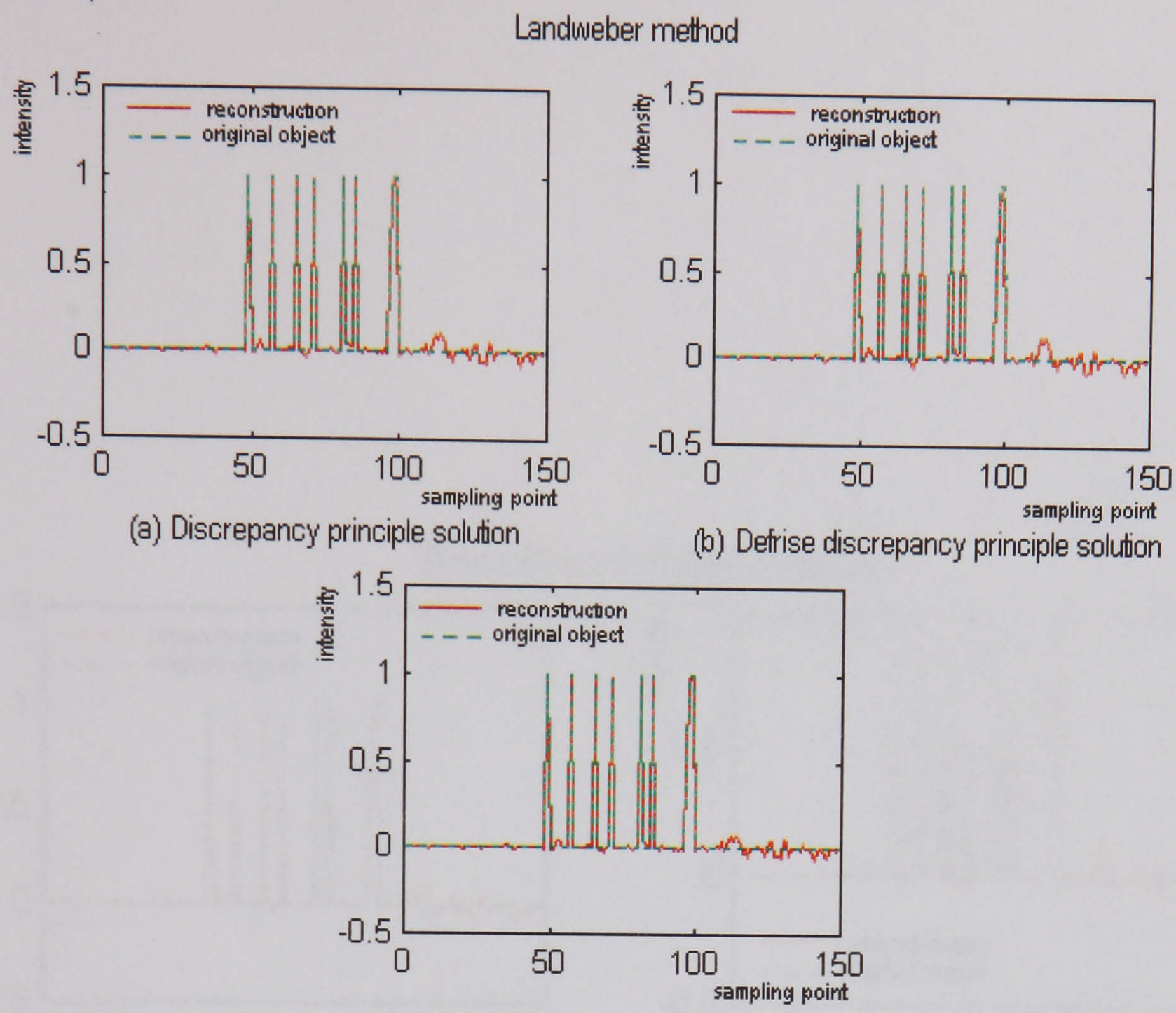


Figure 6.4: Reconstructions (continuous red line) of a *nine point object*, obtained with TSVD regularization method and associated parameter choice criteria, from a *noisy* integral image, versus the real object intensity (dashed green line).

TSVD regularization	Optimal regularization parameter	Relative restoration error (ρ_{rel})	Estimated discrepancy ($\epsilon^2(f, g^{(0)})$)	Estimated object energy (E^2)
Semiconvergence theorem	$J = 131$	5.03%	2.04×10^{-4}	8.97
Discrepancy principle	$J = 117$	10.6%	2.92×10^{-4}	8.88
Discrepancy principle (Defrise)	–	10^{13}%	–	–
Prescribed energy method	$J = 122$	9.47%	2.58×10^{-4}	8.904

Table 6.3: Regularization results for object space intensity reconstruction from a *noisy* integral image containing *nine separate luminous points* using TSVD regularization and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.



(a) Discrepancy principle solution

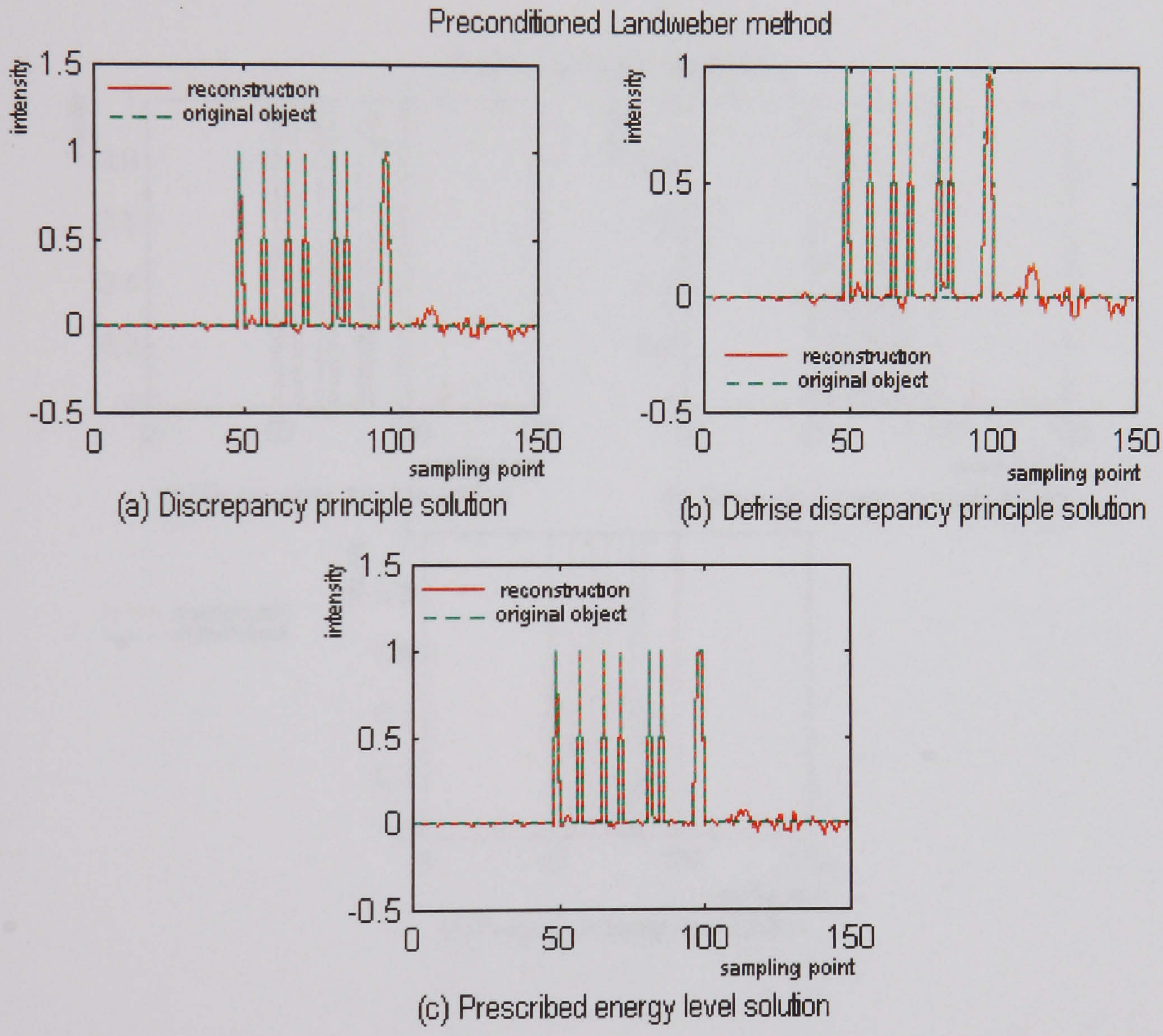
(b) Defrise discrepancy principle solution

(c) Prescribed energy level solution

Figure 6.5: Reconstructions (continuous red line) of a *nine point object*, obtained with Landweber regularization method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

Landweber's method and those derived from it provide a wide range of approximate solutions of the object space reconstruction problem. The results obtained for the reconstruction of the noisy 9-point image are shown in Tables 6.4-6.5 and Figures 6.5-6.9 for each method and its associated parameter choice (stopping) rules.

The relative restoration errors range between 3.58% for the projected Landweber method with prescribed energy level stopping rule and 14.1% for Sanz preconditioned Landweber with Defrise's extension of the discrepancy principle as a stopping rule. From the computational point of view, it can be noticed that, for Landweber, projected Landweber, and Piana preconditioned projected Landweber methods, the number of iterations required is high. Among these more intensive methods, projected Landweber provides the lowest errors for each of the stopping rules. Using a Sanz preconditioner dramatically improves the computational aspect, the necessary number of iterations reducing by more than 10 times, from 700-1200 to 53-113. The errors are moderate for the Sanz preconditioned Landweber (8.58%-14.1%) and



(a) Discrepancy principle solution

(b) Defrise discrepancy principle solution

(c) Prescribed energy level solution

Figure 6.6: Reconstructions (continuous red line) of a *nine point object*, obtained with Preconditioned Landweber regularization (Sanz, $\gamma = 0.1$) method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

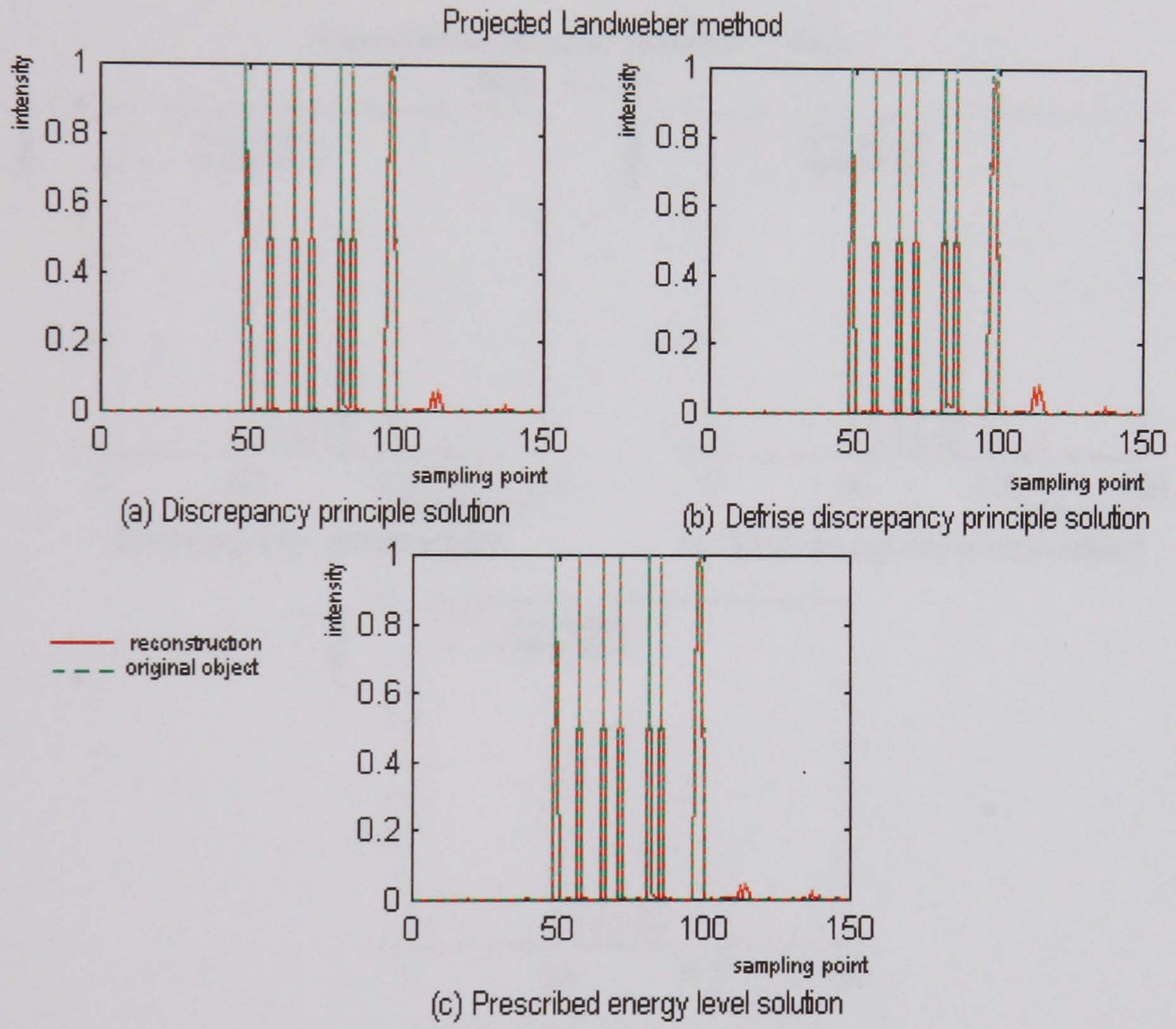


Figure 6.7: Reconstructions (continuous red line) of a *nine point object*, obtained with Projected Landweber regularization method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

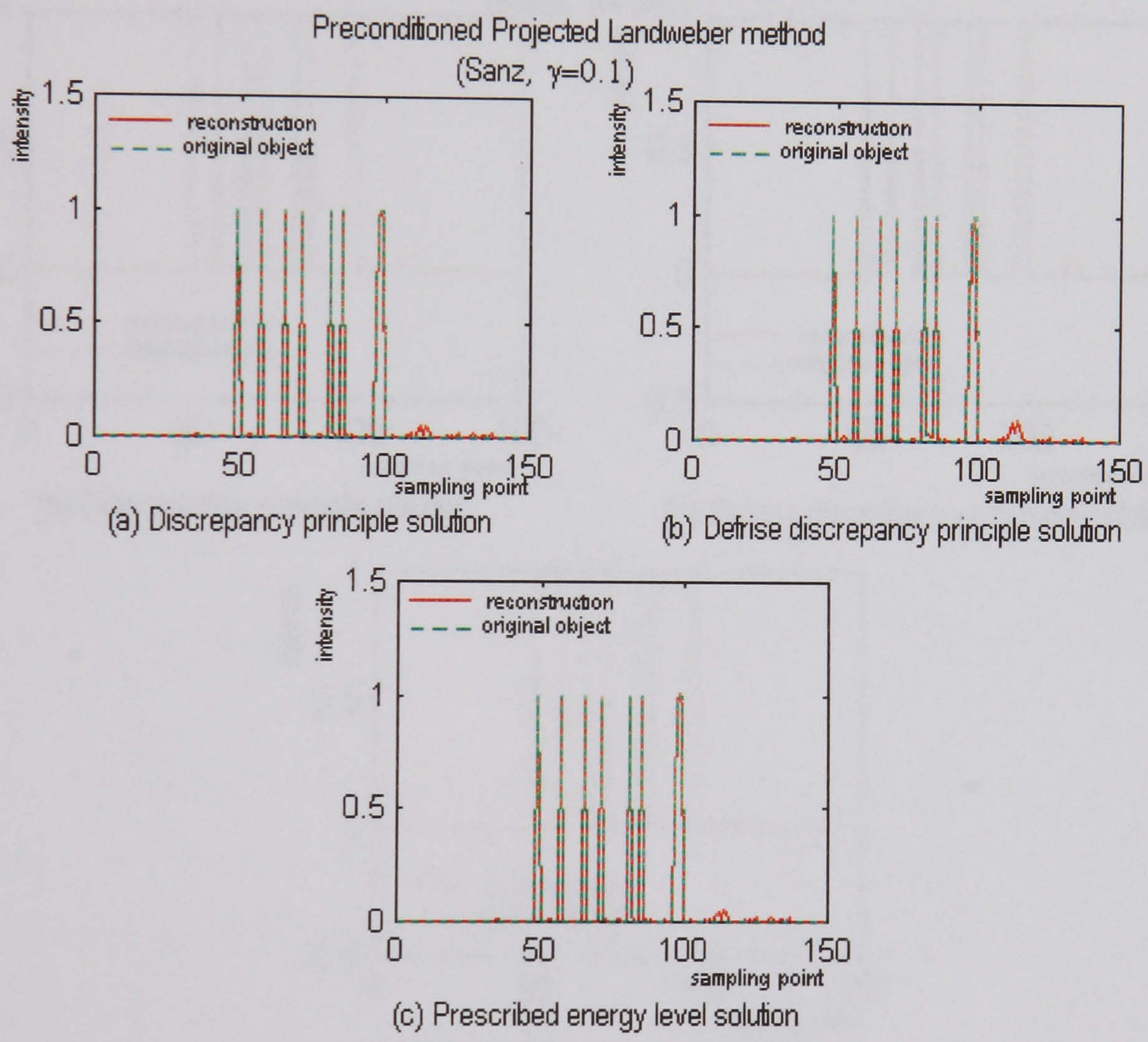


Figure 6.8: Reconstructions (continuous red line) of a *nine point object*, obtained with Preconditioned Projected Landweber regularization (Sanz, $\gamma = 0.1$) method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

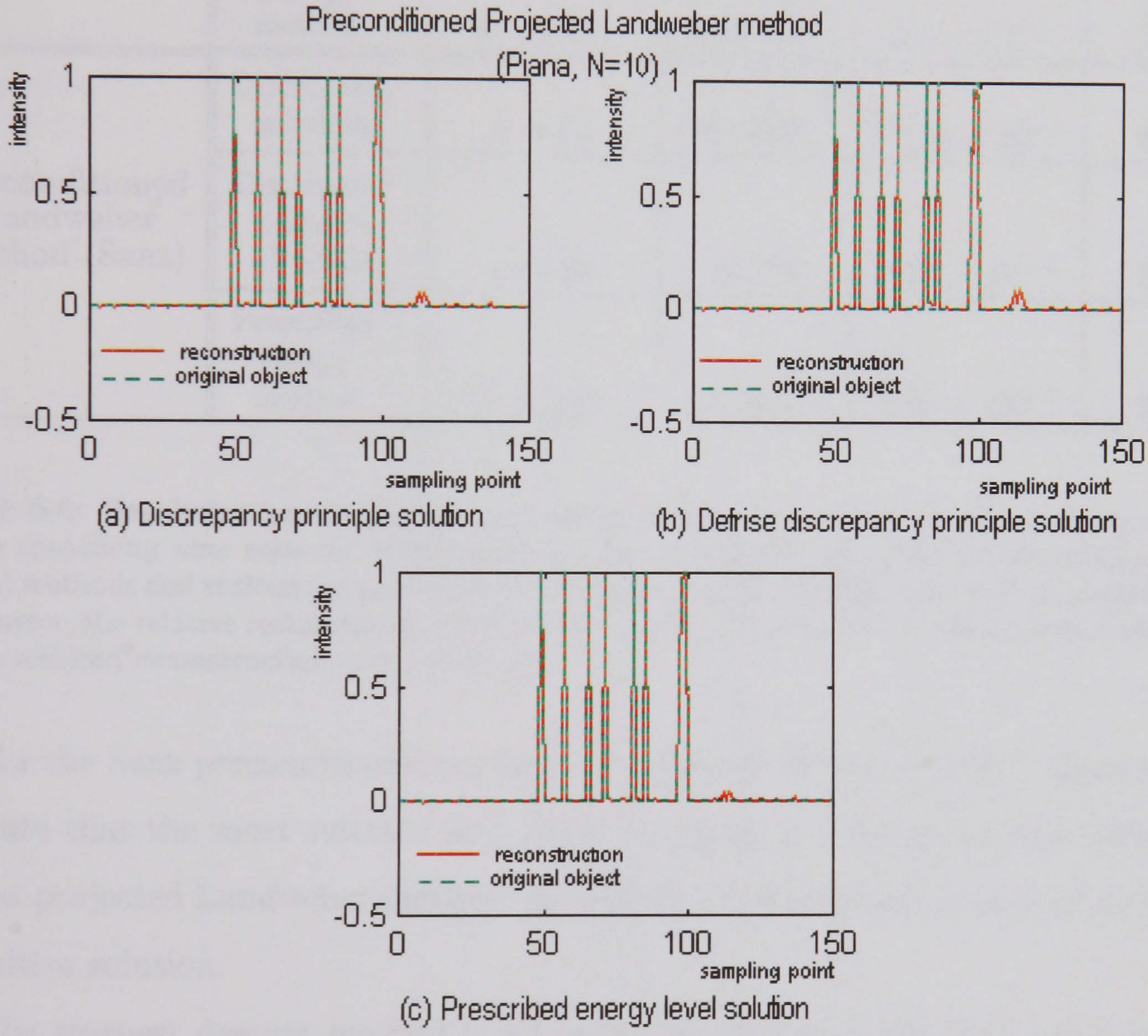


Figure 6.9: Reconstructions (continuous red line) of a *nine point object*, obtained with Preconditioned Projected Landweber regularization (Piana, N=10) method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

Method	Stopping rule	Optimal regularization parameter k	Relative restoration error ρ_{rel}	Estimated discrepancy $\epsilon^2(f,g)$	Estimated object energy E^2
Landweber method	Discrepancy principle	$k = 1240$	10.81%	3.876×10^{-4}	8.57
	Discrepancy principle (Defrise)	$k = 941$	12.90%	5.16×10^{-4}	8.44
	Prescribed energy method	$k = 1727$	8.64%	2.9×10^{-4}	8.70
Preconditioned Landweber method (Sanz)	Discrepancy principle	$k = 81$	10.76%	3.871×10^{-4}	8.57
	Discrepancy principle (Defrise)	$k = 53$	14.1%	6.25×10^{-4}	8.32
	Prescribed energy method	$k = 113$	8.58%	2.89×10^{-4}	8.70

Table 6.4: Regularization results for object space intensity reconstruction from a *noisy* integral image containing *nine separate luminous points* using Landweber and Preconditioned Landweber (Sanz) methods and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

low for the Sanz preconditioned projected Landweber (4.04%-7.06%). These results indicate that the most suitable and efficient approach is using the Sanz preconditioned projected Landweber method. In addition, it has the advantage of providing a positive solution.

The steepest descent method is a Landweber regularization with optimised relaxation parameter. The characteristics of reconstructions via steepest descent are shown in Table 6.6. The semiconvergence theorem gives a very high iteration number for the optimal solution (3540), compared to the approximation criteria, with very little gain in accuracy (relative error 5.01% compared to 5.74% obtained with the prescribed energy method). Miller's method provides an optimal iteration number $k = (E/\epsilon)^2 \simeq 2.3 \times 10^4$, which is similarly highly overestimated. Moderate restoration errors are obtained with the discrepancy principle (10.8%) whilst the iteration number is 311, much lower than that indicated by Miller's method or by the semiconvergence theorem. The prescribed energy level stopping rule gives a solu-

Method	Stopping rule	Optimal regularization parameter k	Relative restoration error ρ_{rel}	Estimated discrepancy $\epsilon^2(f,g)$	Estimated object energy E^2
Projected Landweber method	Discrepancy principle	$k = 862$	4.55%	3.875×10^{-4}	8.62
	Discrepancy principle (Defrise)	$k = 721$	6.24%	5.16×10^{-4}	8.48
	Prescribed energy method	$k = 983$	3.58%	3.3×10^{-4}	8.70
Preconditioned Projected Landweber method (Sanz)	Discrepancy principle	$k = 73$	4.31%	3.83×10^{-4}	8.67
	Discrepancy principle (Defrise)	$k = 54$	7.06%	6.26×10^{-4}	8.46
	Prescribed energy method	$k = 76$	4.04%	3.65×10^{-4}	8.70
Preconditioned Projected Landweber method (Piana), $N = 10$	Discrepancy principle	$k = 867$	5.20%	3.85×10^{-4}	8.58
	Discrepancy principle (Defrise)	$k = 728$	6.99%	5.15×10^{-4}	8.44
	Prescribed energy method	$k = 1043$	3.76%	3.05×10^{-4}	8.70

Table 6.5: Regularization results for object space intensity reconstruction from a *noisy* integral image containing *nine separate luminous points* using Projected Landweber methods and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

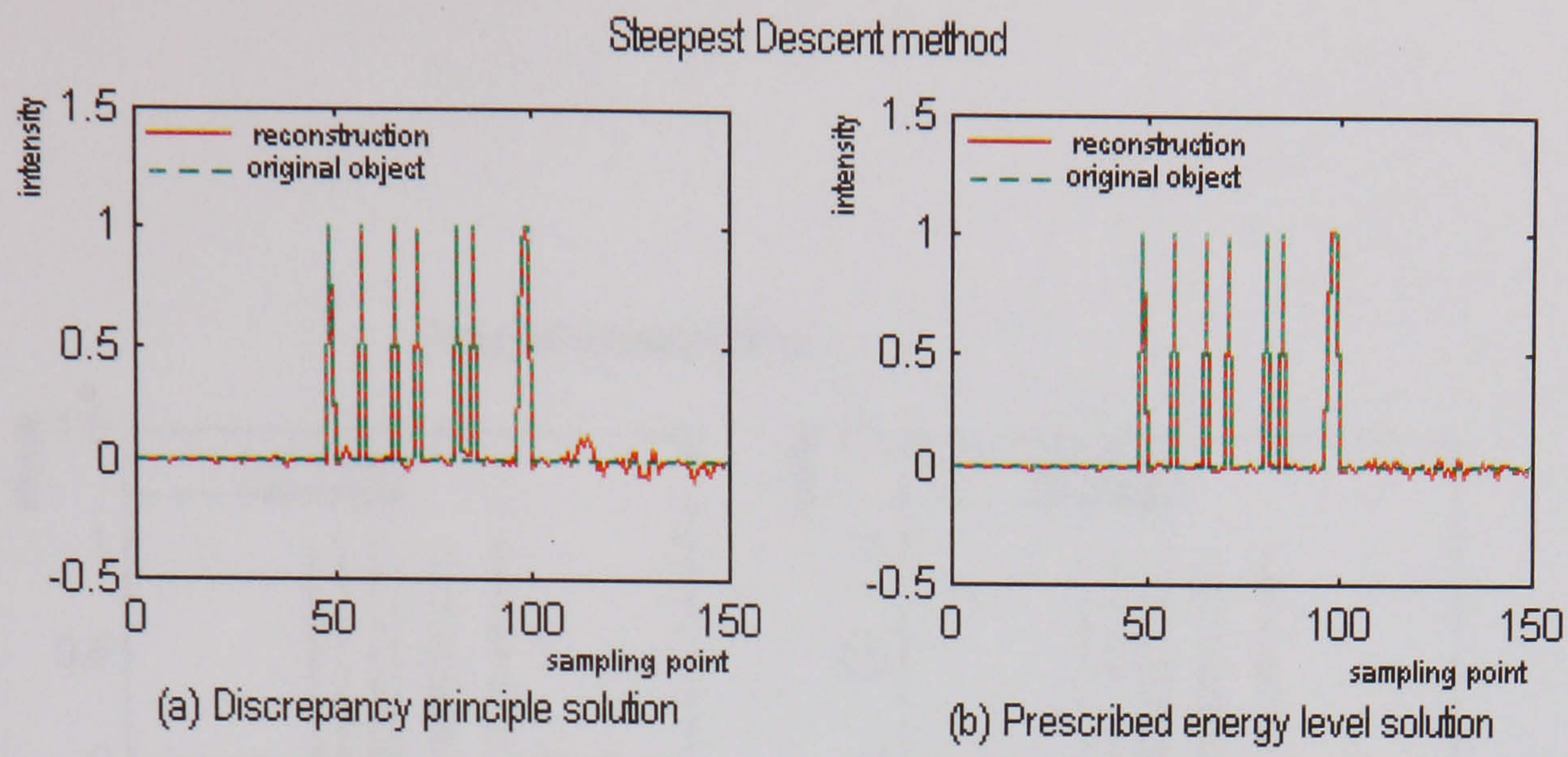


Figure 6.10: Reconstructions (continuous red line) of a *nine point object*, obtained with Steepest Descent regularization method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

Steepest descent method	Optimal regularization parameter	Relative restoration error (ρ_{rel})	Estimated discrepancy ($\epsilon^2(f, g^{(0)})$)	Estimated object energy (E^2)
Semiconvergence theorem	$k = 3540$	5.01%	1.9×10^{-4}	8.93
Discrepancy principle	$k = 311$	10.8%	3.87×10^{-4}	8.58
Prescribed energy method	$k = 994$	5.74%	2.12×10^{-4}	8.90

Table 6.6: Regularization results for object space intensity reconstruction from a *noisy* integral image containing *nine separate luminous points* using Steepest descent method and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

tion in 911 iterations with a good approximation (5.74%). As a general observation, the Steepest Descent method is an optimisation of the Landweber method that has slower convergence and is more computationally intensive than the preconditioned versions of the Landweber method. The regularization results are presented in Table 6.6 and Figure 6.10.

The results obtained with the conjugate gradient method for the 9-point object reconstruction from a noisy image are shown in Table 6.7 and Figure 6.11. The method is extremely fast and reliable. The errors obtained are moderate, ranging between 5% and 9.24%.

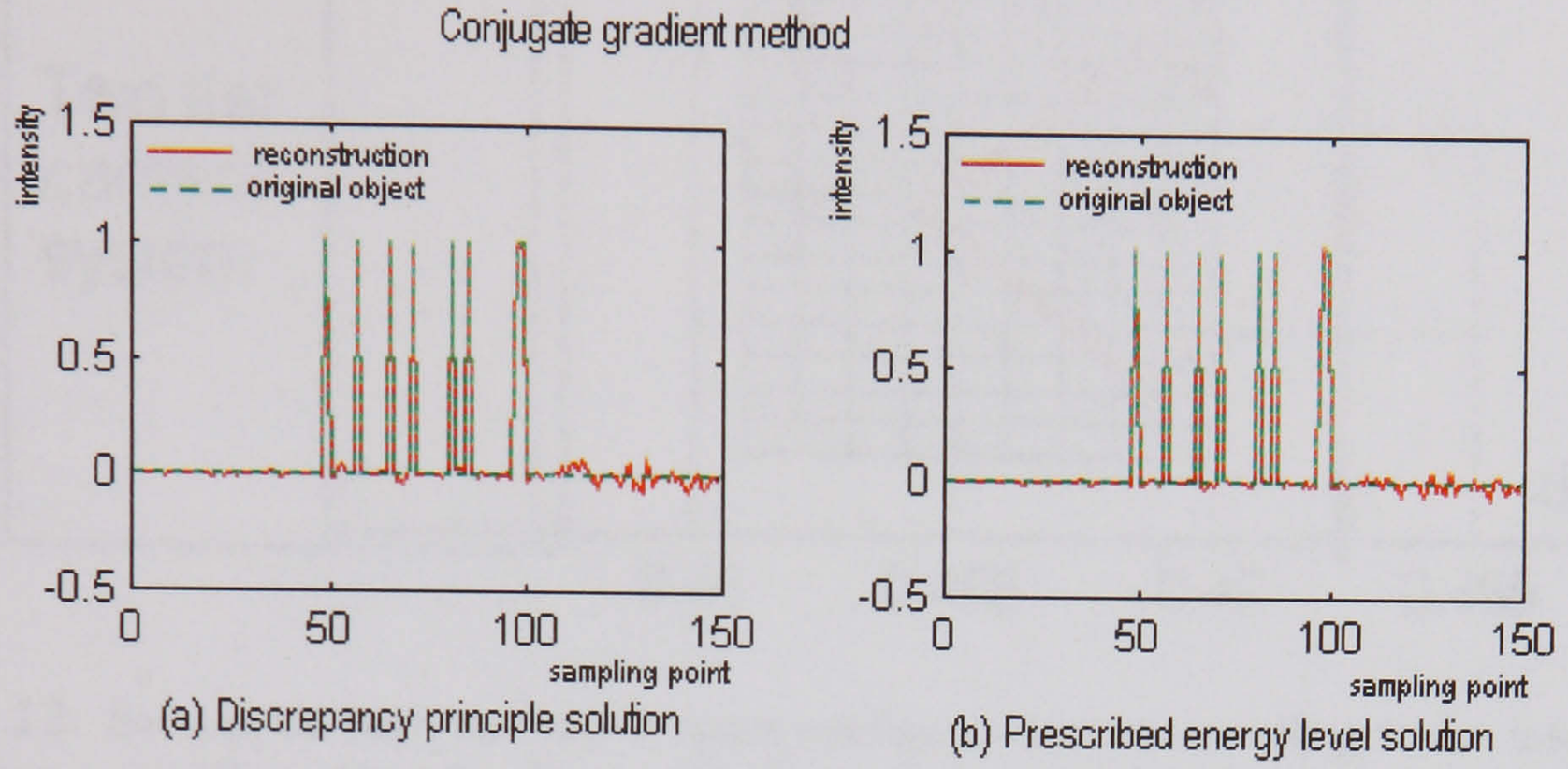


Figure 6.11: Reconstructions (continuous red line) of a *nine point object*, obtained with Conjugate Gradient regularization method and associated stopping rules, from a *noisy* integral image, versus the real object intensity (dashed green line).

Conjugate gradient method	Optimal regularization parameter	Relative restoration error (ρ_{rel})	Estimated discrepancy ($\epsilon^2(f, g^{(0)})$)	Estimated object energy (E^2)
Semiconvergence theorem	$k = 100$	5.03%	2.04×10^{-4}	8.97
Discrepancy principle	$k = 32$	9.24%	3.46×10^{-4}	8.77
Prescribed energy method	$k = 38$	7.01%	2.36×10^{-4}	8.903

Table 6.7: Regularization results for object space intensity reconstruction from a *noisy* integral image containing *nine separate luminous points* using Conjugate Gradient method and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

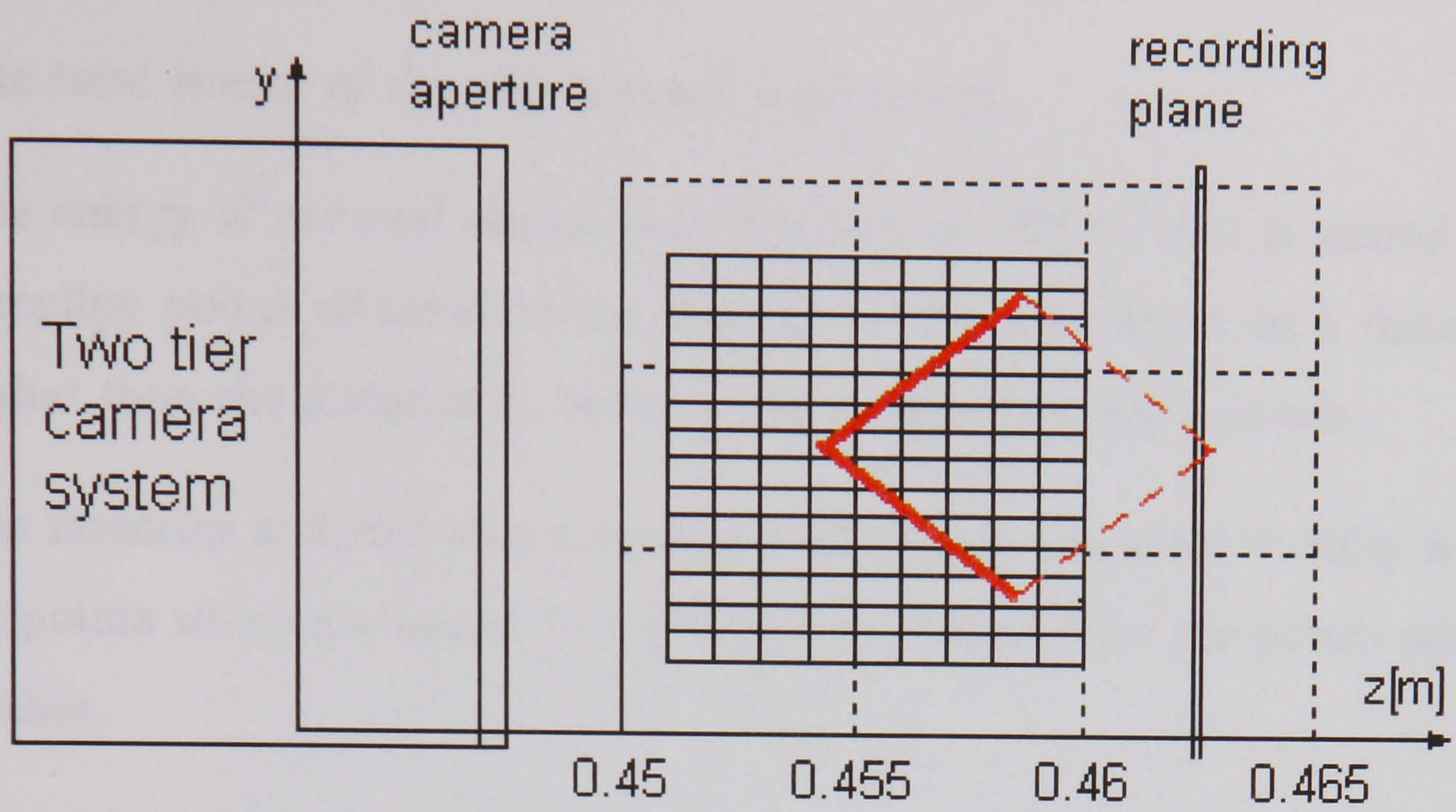


Figure 6.12: Section through the object space configuration corresponding to the integral image of an opaque cube (Example 6.3). The visible faces of the cube are shown in continuous line; the obscured faces are represented by the dashed line.

6.3 Example 2. Reconstruction of the visible faces of an opaque cube

The second computer generated integral image analysed contains an opaque cube which has a couple of faces parallel to the Oyz plane and the other four faces oriented at $\pm 45^\circ$ with regard to the Oz-axis. Its intersections with $x = x_0$ planes are similar to each other and look like the one shown in Figure 6.12, therefore the recovery of one x-section automatically leads to the recovery of the whole object. For simplicity and clarity reasons, planar sections parallel to the Oyz-plane of the object space reconstruction are shown in Figures 6.16, 6.27-6.35, which contain both the object reconstruction and the original object contour for comparison. The integral image is shown in Figure 8.8, Appendix C.

The image information was analysed using a 150-point sampling grid in the object space. The grid does not contain the points of the original object, so it is not possible to have an exact 'object model' to compare with the reconstruction and thereby obtain the relative restoration error. It is plain that assessing the image restoration performance is affected by the 'object model' chosen as best match to the original object at the points of the sampling grid.

The 'object model' chosen to represent the imaged object - namely, the visible

part of the cube - has been constructed making the following assumptions:

1. The total energy of the object space is preserved.

2. The energy of the real object (which is known: $E^2 = 161$) is shared by the sampling points situated in the vicinity of the real object at a distance no higher than the distance d_z between two consecutive depth levels.

3. The intensity assigned to a sampling point within the given vicinity is higher for points which are nearer to the object and smaller for the points which are farther.

In order to fulfill these conditions, the visible (recorded) part of the object has been divided into the two faces - F_1 and F_2, each of which has an energy equal to half of the whole energy. Two weights $w_i^{F_1}$ and $w_i^{F_2}$ have been assigned to each sampling point i:

$$w_i^{F_k} = \begin{cases} 0\,, & \text{if } d_i^{F_k} > d_z \\ \dfrac{\sum_j d_j^{F_k}}{d_i^{F_k} + \sum_j d_j^{F_k}}\,, & \text{if } 0 \le d_i^{F_k} \le d_z\,, \end{cases} \tag{6.2}$$

where $d_i^{F_k}$ is the distance from the point i to the real object F_k, $k = 1, 2$. The weight $w_i^{F_k}$ is 1 for a point situated on the object F_k and less than 1 for points which do not belong to the face. The further the point is from the object, the lower its weight is. The intensity value assigned to a sampling point is:

$$\tilde{f}_{F_k}^{(0)}(i) = w_i^{F_k} \frac{E^2/4}{\sum_j w_j^{F_k}}\,, \tag{6.3}$$

where $(E^2/4)/\sum_j w_j^{F_k}$ is the weighted mean of the energy of the real object face.

The resulting vector $\tilde{f}^{(0)} = \tilde{f}_{F_1}^{(0)} + \tilde{f}_{F_2}^{(0)}$ is adopted as the *expected intensity picture (map)* of the object space (Figure 6.13). It does not represent the real object, but an energy conservative interpolation of it at the points of the sampling grid.

The reconstructions f of the object space obtained from the integral image via various regularization methods are compared with this expected result to obtain a *relative pseudo-restoration error*:

$$\tilde{\rho}_{rel} = \frac{\|f - \tilde{f}^{(0)}\|}{\|\tilde{f}^{(0)}\|}\,. \tag{6.4}$$

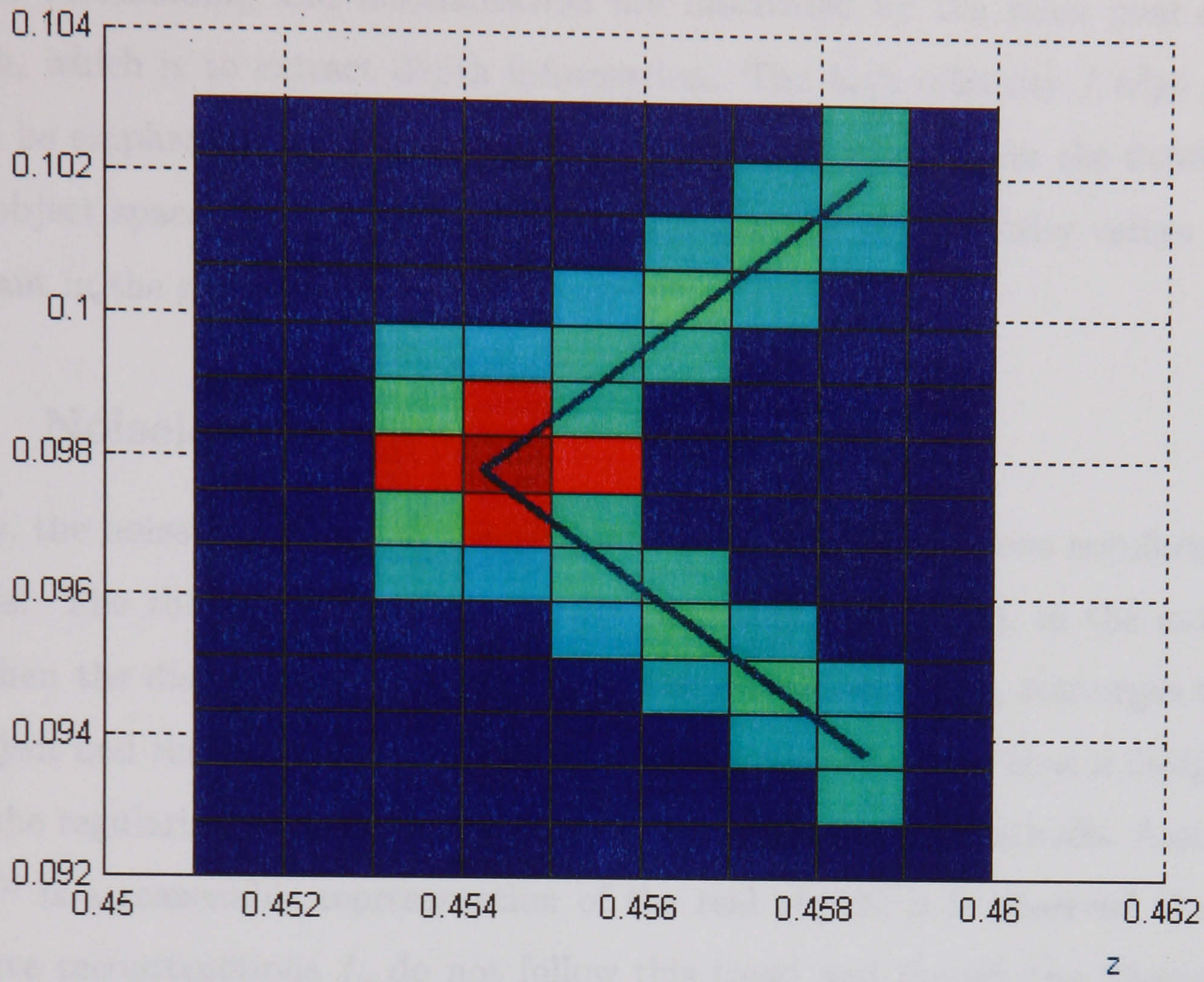


Figure 6.13: Expected intensity map of the object space in the given sampling grid.

The relative pseudo-restoration error is a measure of the performance of the regularization methods providing reconstructions of the object space for the given image. However, its value is more important when comparing two methods than *per se*, as interpolation acts effectively as noise and affects both the expected object space model and the reconstruction.

In order to overcome the difficulty produced by interpolation in assessing the performance of each method, additional processing of the results has been introduced:

1. *'2 out of 3 bins' automatic thresholding rule.* A uniform thresholding rule has been used: the intensity values of the reconstruction vector f have been organised into a 3-equidistant bin histogram, and the values contained in the two lower bins have been set to 0, that is considered background.

2. *Normalisation of the thresholded result.* The thresholded regularization solution has been normalised, i.e. divided by its maximum value such that all the values are between 0 and 1.

Both thresholding and normalisation are motivated by the main goal of this research, which is to extract depth information. The high intensity / edge points have to be emphasised in order to recover the contours which make the depth map of the object space. Therefore, the absolute accuracy of the intensity values is less important in the given context.

6.3.1 Noiseless image

Initially, the noise-free image of the cube was analysed with various regularization methods. The theoretical analysis already presented assert that, in the noiseless case, when the discrepancy ϵ approaches 0, the reconstruction f_μ converges to the true object and the relative restoration error converges to 0. Note that μ designates any of the regularization parameters used by the regularization methods. Assuming that $\tilde{f}^{(0)}$ is a reasonable representation of the real object, it is observed that the successive reconstructions f_μ do not follow this trend and though the discrepancy descends to 0, the relative pseudo-restoration error first decreases to a minimum, and then increases, as depicted in Figure 6.14 for the TSVD reconstruction. At the same time, the reconstruction f_μ tends to take the shape of the real object for values of the regularization parameter on the descending slope of the pseudo-restoration error, however when the discrepancy is almost 0, any relationship to the true object is lost. This statement is proven by Figures 6.15(a)-(b) which present the TSVD reconstruction for $J = J_{opt} = 49$ for which the discrepancy $\epsilon^2 = 2.07$ and the TSVD reconstruction for $J = 135$ for which the discrepancy is $\epsilon^2 = 0.67$, so much smaller. To sum up, the regularization process displays only a semiconvergence behaviour even in the case of the noiseless image. This is due to the noise-like effect of the interpolation. Therefore, it is necessary to consider as optimal the reconstructions for which the relative pseudo-restoration error is a minimum.

The assessment of these optimal reconstructions are presented in Table 6.8. Comparing these results with those obtained from the noiseless nine point picture (Table 6.1), it is noticed that the minimum relative pseudo-restoration error takes much higher values, ranging between 77.86% and 82.78%. The discrepancy corresponding to the optimal reconstructions takes values between 1.72 and 2.64. This figure is an indication on the value of the discrepancy when using the discrepancy principle for

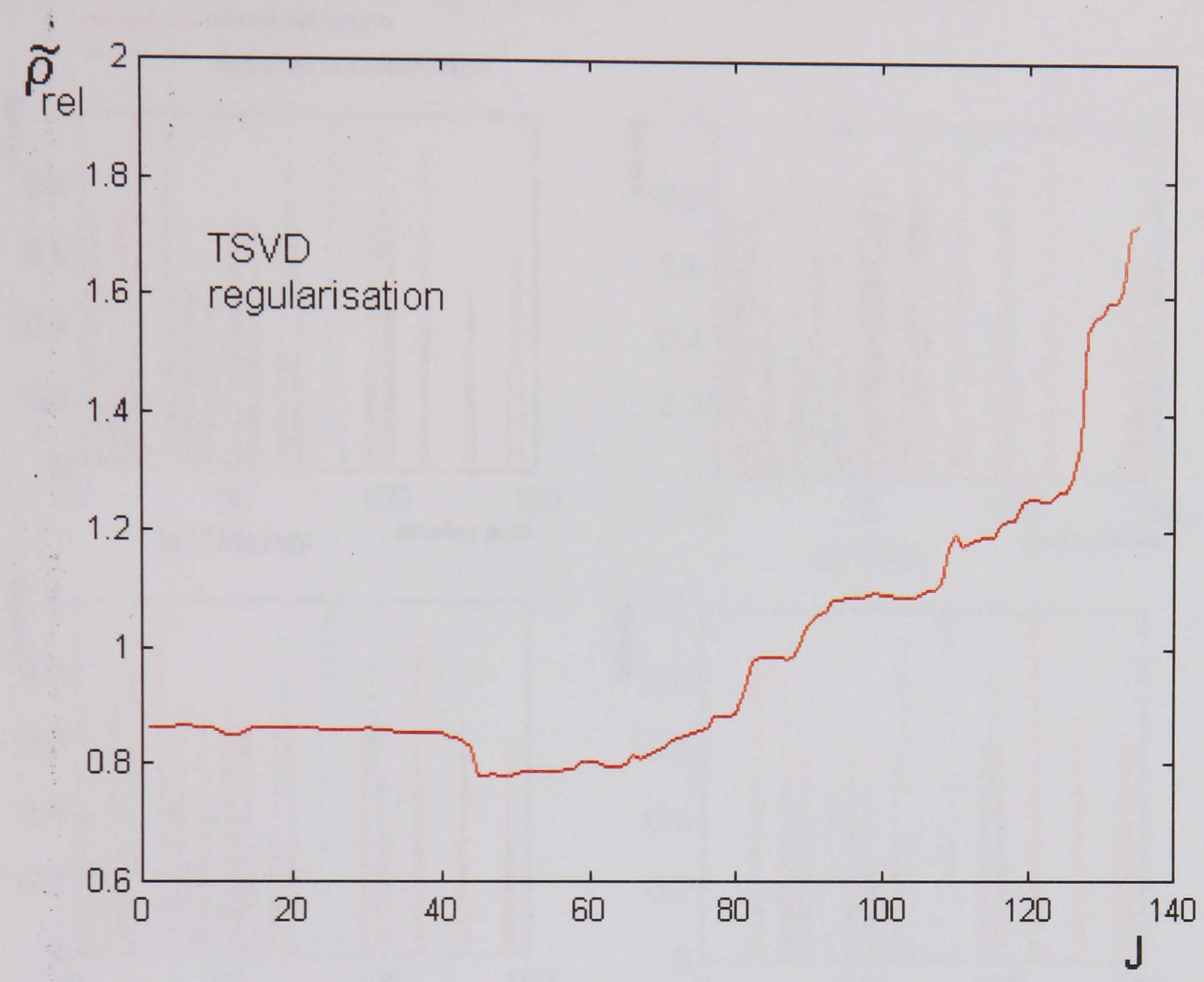


Figure 6.14: The relative restoration error corresponding to the reconstruction of the object space from the noiseless integral image of nine luminous points considered in Example 6.3, using TSVD.

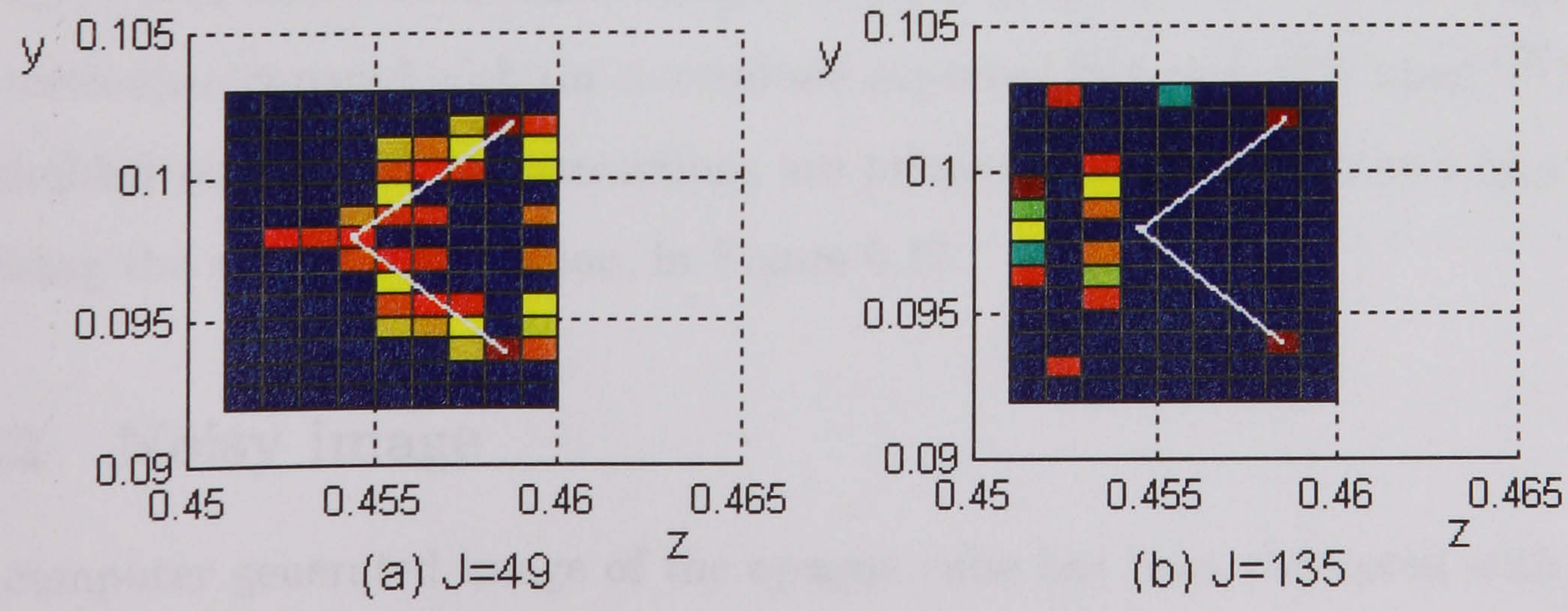


Figure 6.15: Transversal section through the reconstruction of the object space from the noiseless integral image of an opaque cube using TSVD with two values of the regularization parameter: (a) $J = J_{opt} = 49$, (b) $J = 135$. The actual object is depicted by the white continuous line.

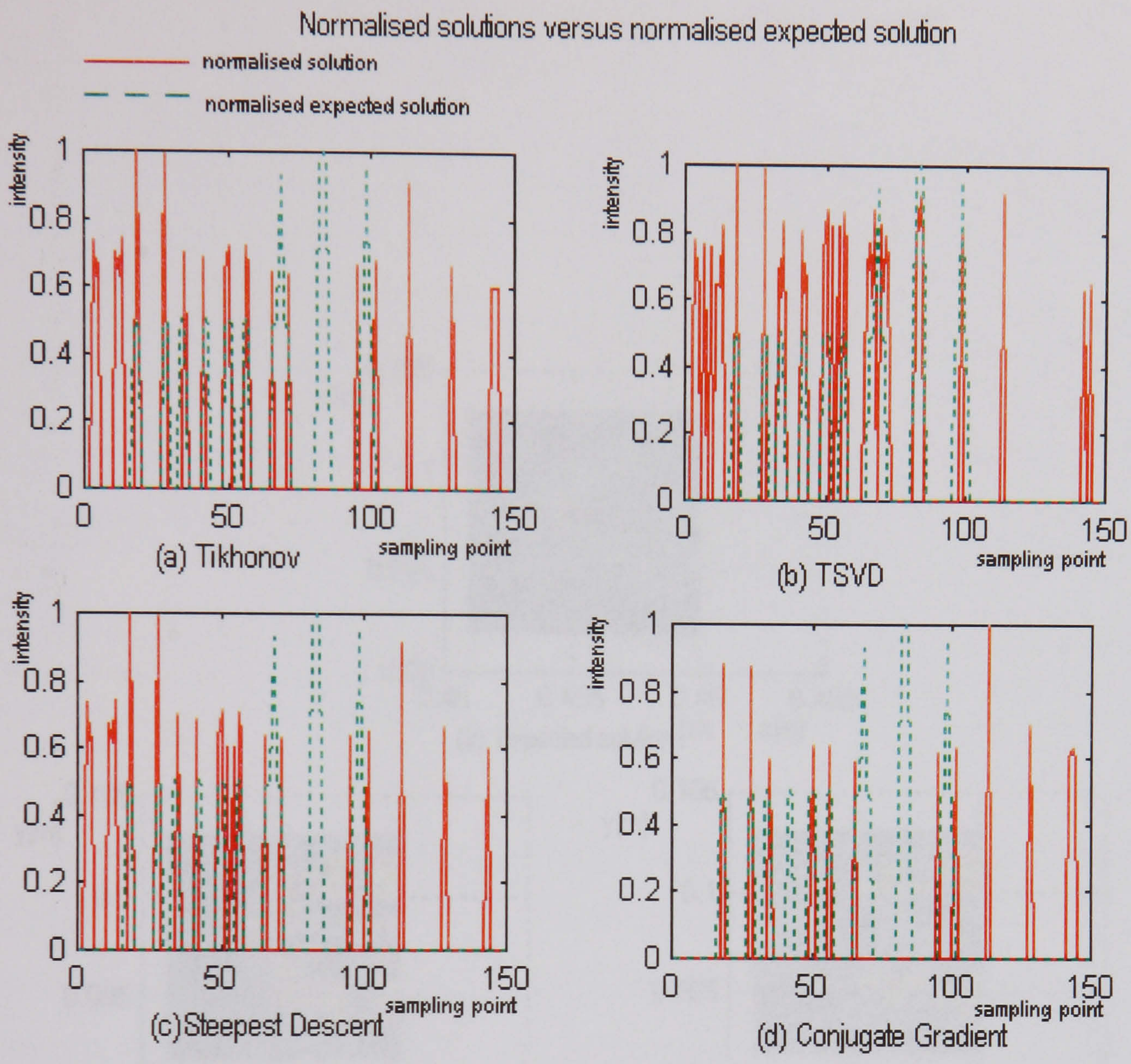


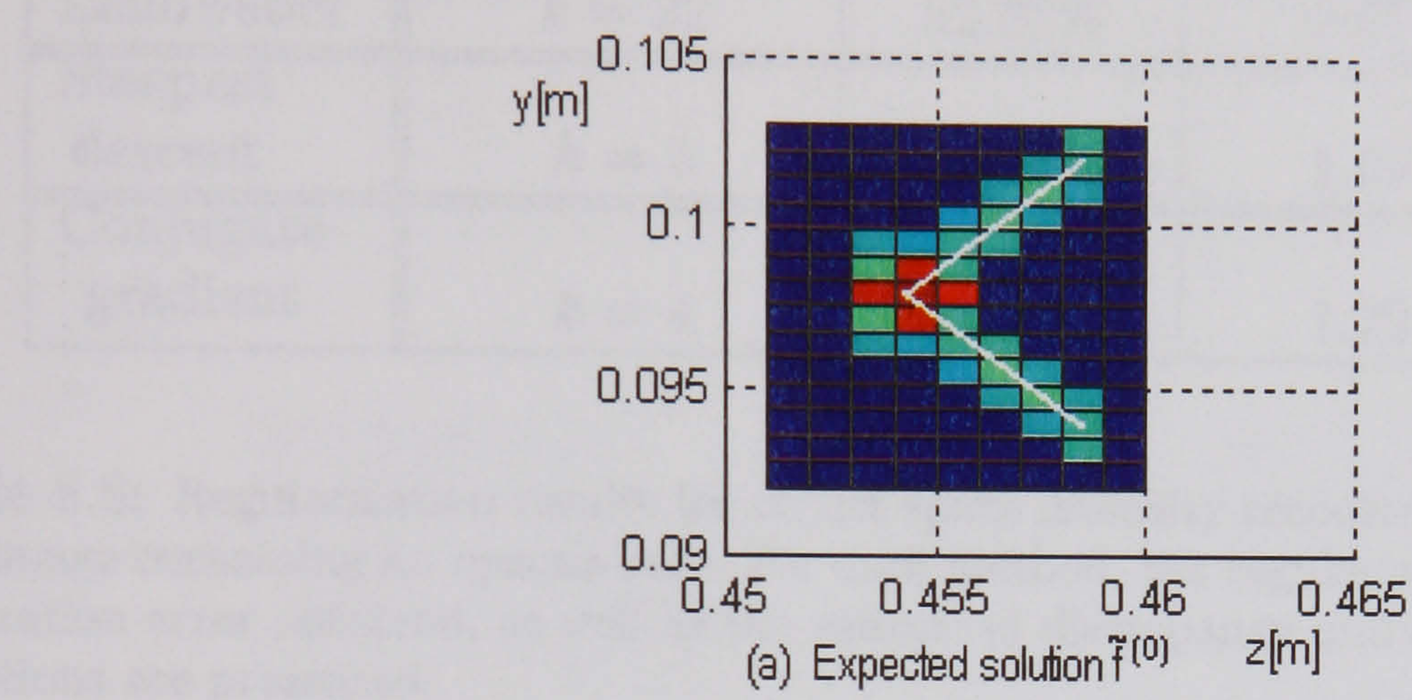
Figure 6.16: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by regularization from a *noiseless* integral image, versus the normalised expected object space intensity (dashed green line).

the analysis of the corresponding noisy picture.

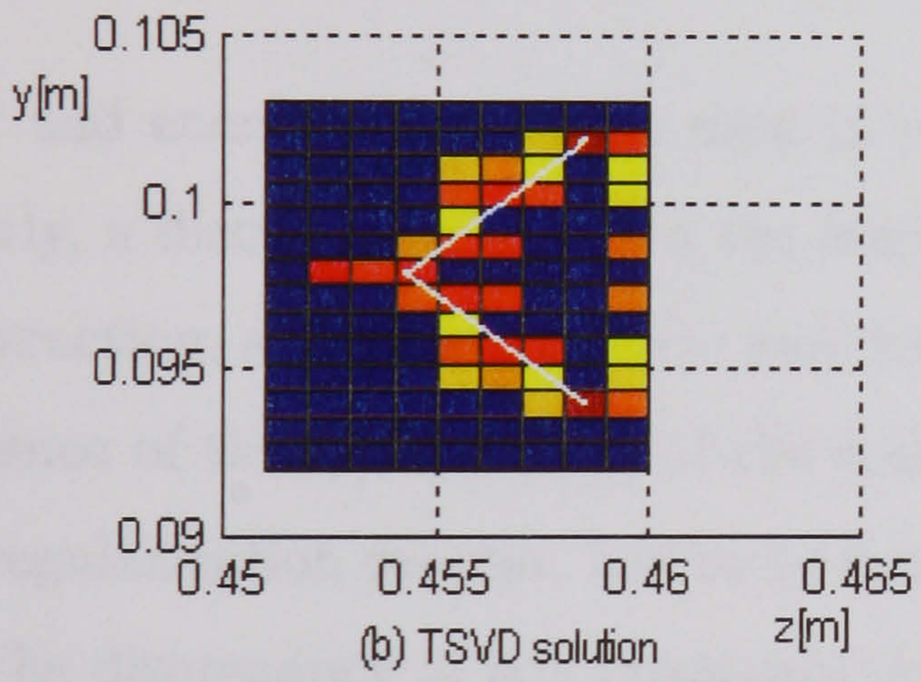
Figure 6.16 shows normalised reconstructions obtained with various regularization methods compared with the normalised expected intensity $\tilde{f}^{(0)}/\max(\tilde{f}^{(0)})$. The thresholded regularized reconstructions are presented using coordinates (y, z), emphasising the spatial distribution, in Figure 6.17.

6.3.2 Noisy image

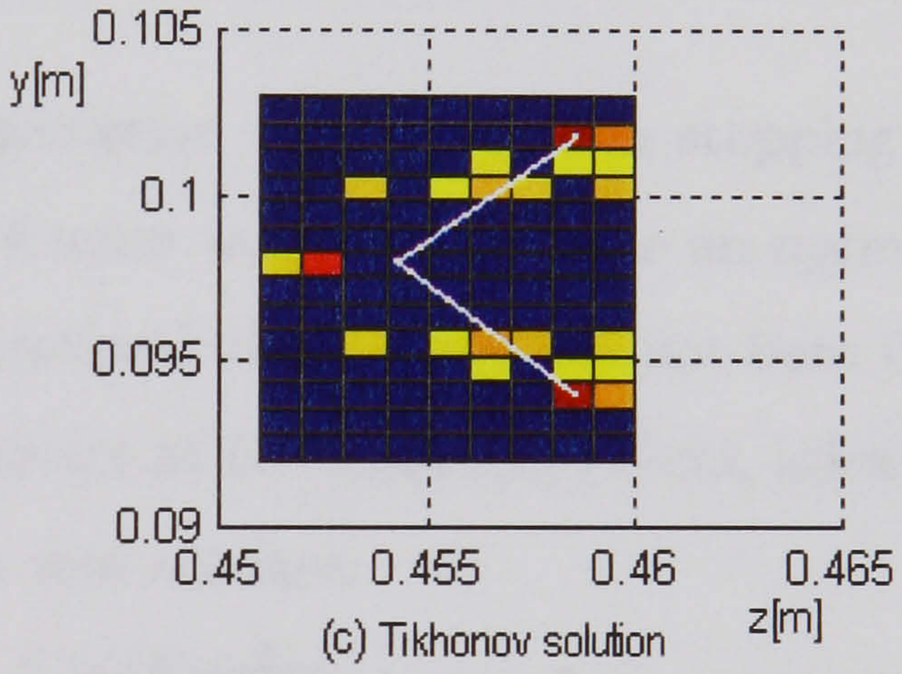
The computer generated image of the opaque cube has been corrupted with white Gaussian noise of mean 0 and variance equal to the square of 1% of the maximum intensity value in the noiseless picture. The results of the reconstructions using various regularization methods are shown in Tables 6.9-6.11 and Figures 6.18-6.35. The main problem posed by regularization is encountered when chosing the discrep-



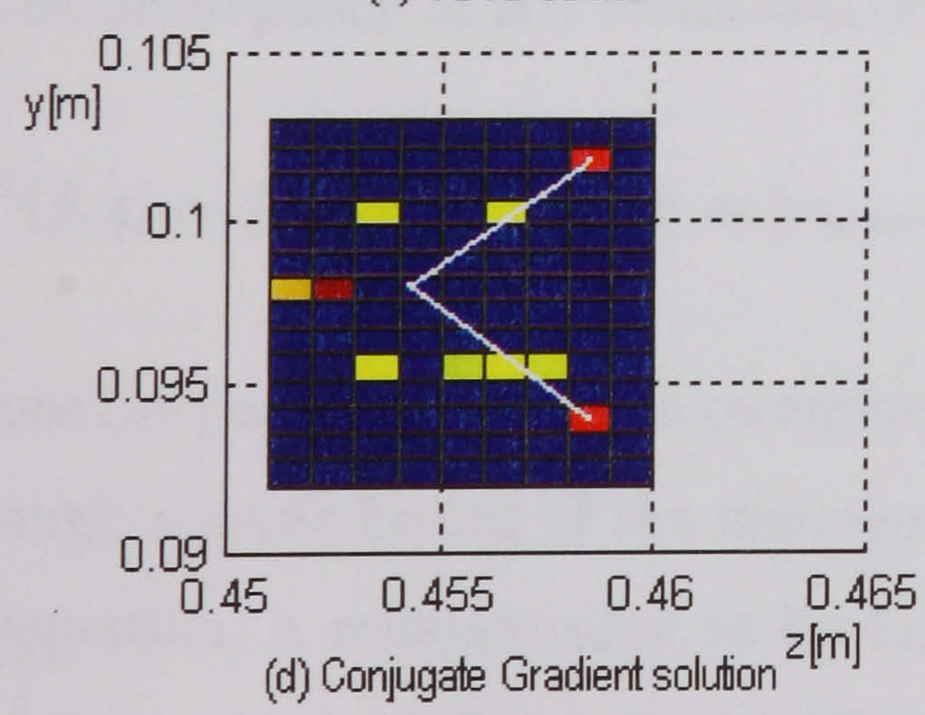
(a) Expected solution $\tilde{f}^{(0)}$



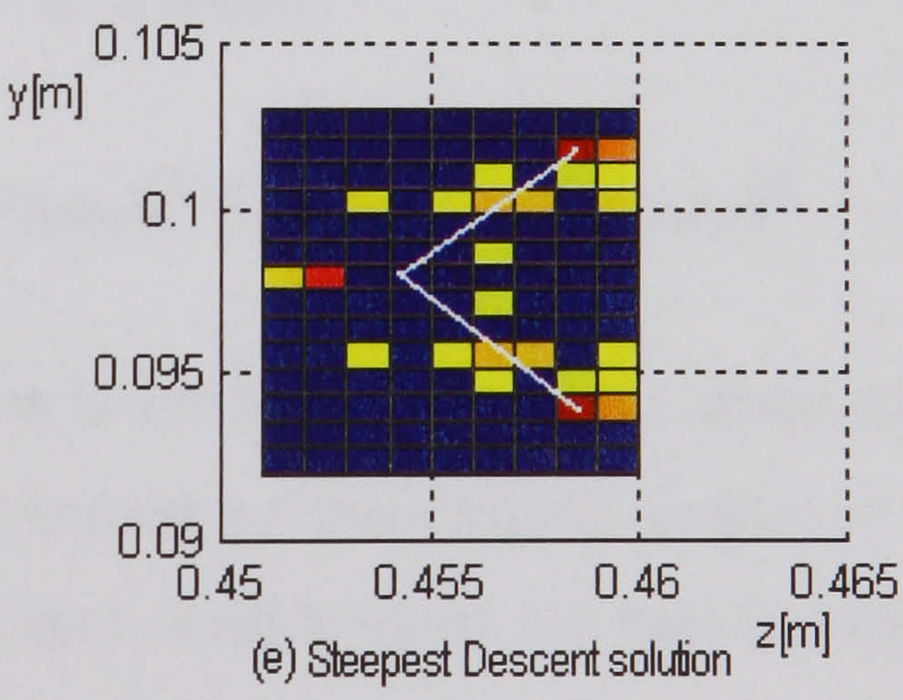
(b) TSVD solution



(c) Tikhonov solution



(d) Conjugate Gradient solution



(e) Steepest Descent solution

Figure 6.17: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by regularization from a *noiseless* integral image.

Method	Regularization parameter	Relative restoration error (ρ_{rel})	Estimated discrepancy ($\epsilon^2(f, g^{(0)})$)	Estimated object energy (E^2)
Tikhonov	$\mu = 0.0225$	82.78%	2.64	229.17
TSVD	$J = 49$	77.86%	2.07	321.78
TSVD with positivity	infeasible	–	–	–
Landweber	$k = 22$	82.21%	1.97	295.5
Steepest descent	$k = 5$	82.29%	1.99	300.14
Conjugate gradient	$k = 4$	82.31%	1.72	337.70

Table 6.8: Regularization results for object space intensity reconstruction from a *noiseless* integral image containing an *opaque cube*. For each method, the regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

ancy and energy values to be used in conjunction with the known stopping rules. Clearly, a discrepancy equal to the level of noise would not provide an optimal reconstruction, as in the noise-free case the optimal discrepancy had not been 0. The influence of the interpolation of the real object at the sampling points, inherent to the regularization process, has to be taken into account.

The discrepancy of any reconstruction f is therefore:

$$\epsilon^2(f, g) = \|Af - g\|^2 = \|\nu\| = \|\nu_{noise} + \nu_{interp}\|^2 \leq \|\nu_{noise}\|^2 + \|\nu_{interp}\|^2 , \quad (6.5)$$

because the perturbation in the reconstruction is due both to noise and interpolation. Although a lower bound of the discrepancy cannot be theoretically predicted from this equation, a reliable value to be taken into consideration for the discrepancy principle is obtained from the noiseless case. From Table 6.8, it is noticed that the discrepancy corresponding to the optimal reconstructions has a value around 2. This is considered as a reference value for the perturbation due to interpolation. The perturbation due to the noise added to the image is given by:

$$\|\nu_{noise}\|^2 = 0.0192 . \quad (6.6)$$

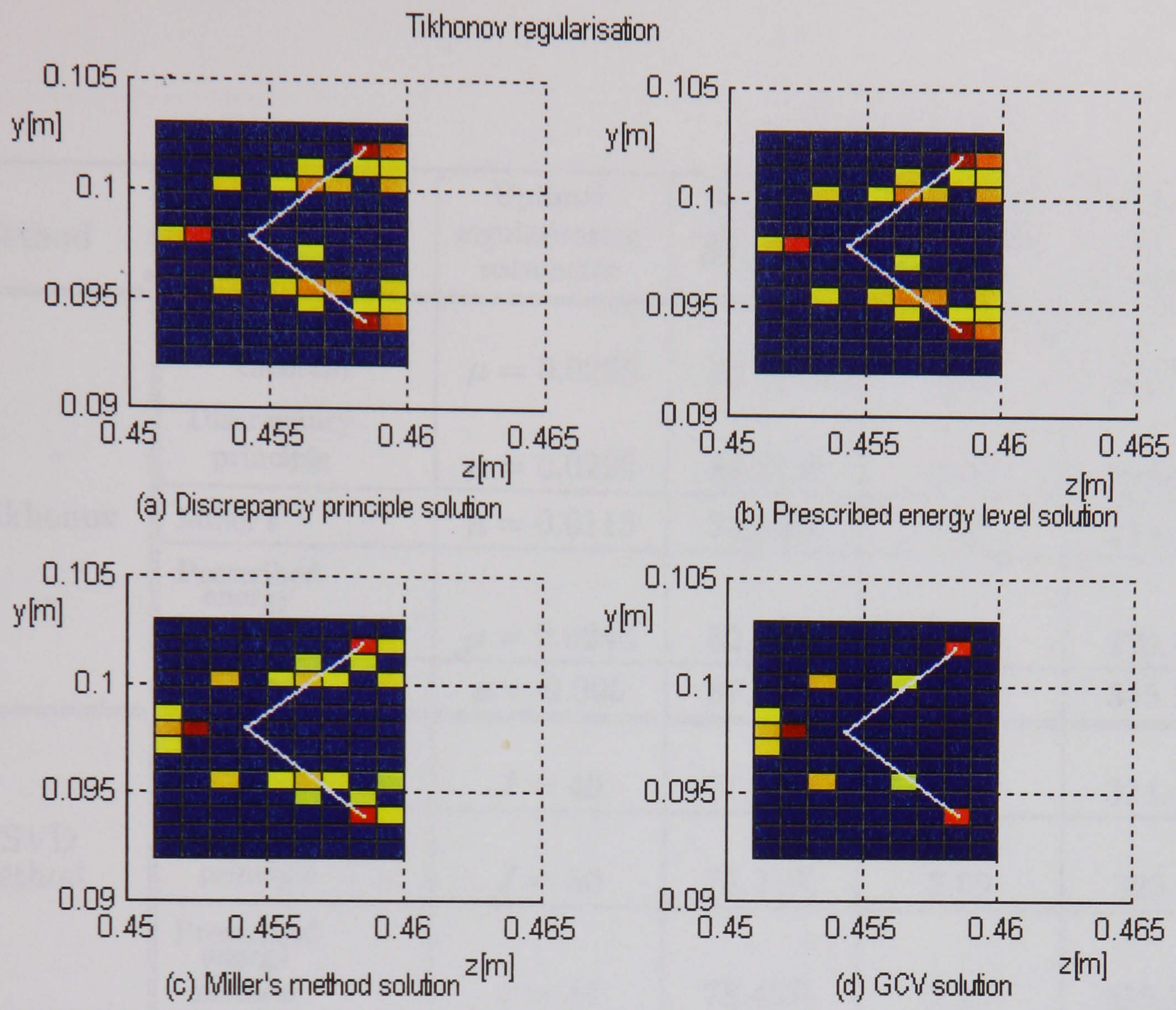


Figure 6.18: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Tikhonov regularization and associated parameter choice criteria from a *noisy* integral image.

For the discrepancy principle, a standard value $\epsilon^2 = 2.0192$ was considered.

For the same reasons, a reference value of the energy to be prescribed to the optimal object is not the real energy value (161), but the value which averages the energy values of the optimal objects obtained by regularization from the noiseless picture: $E^2 = 300$.

These reference values have been confirmed by the semiconvergence theorem results for each of the regularization methods, as shown in Tables 6.9-6.11.

The values of the relative pseudo-restoration errors are high, ranging between 77% and 83%, and the quality differences between results obtained with different methods are not significant. In spite of these high error figures, the shape of the reconstructed object is correctly reproduced by the high intensity values in each reconstruction. This is better seen when one considers the thresholded version of the reconstructions (Figures 6.18-6.26).

Figures 6.27-6.35 comprise comparisons between the normalised expected inten-

Method	Stopping rule	Optimal regularization parameter	Relative restoration error ρ_{rel}	Estimated discrepancy $\epsilon^2(f,g)$	Estimated object energy E^2
Tikhonov	Semiconvergence theorem	$\mu = 0.0225$	82.70%	2.69	229.67
	Discrepancy principle	$\mu = 0.0225$	82.70%	2.69	229.67
	Miller's	$\mu = 0.0113$	83.55%	1.91	277.56
	Prescribed energy method	$\mu = 0.0246$	82.71%	2.69	229.74
	GCV	$\mu = 0.005$	87.27%	1.40	345.10
TSVD method	Semiconvergence theorem	$J = 49$	77.80%	2.10	324.40
	Discrepancy principle	$J = 50$	78.21%	2.09	325.6
	Prescribed energy method	$J = 47$	78.45%	2.26	315.39
Steepest descent method	Semiconvergence theorem	$k = 5$	82.20%	2.04	301.15
	Discrepancy principle	$k = 6$	82.22%	1.91	288.79
	Prescribed energy method	$k = 5$	82.20%	2.04	301.15
Conjugate gradient method	Semiconvergence theorem	$k = 4$	82.25%	1.75	340.85
	Discrepancy principle	$k = 4$	82.25%	1.75	340.85
	Prescribed energy method	$k = 4$	82.25%	1.75	340.85

Table 6.9: Regularization results for object space intensity reconstruction from a *noisy* integral image containing an *opaque cube* using Tikhonov, TSVD, Steepest Descent, and Conjugate Gradient methods and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

Method	Stopping rule	Optimal regularization parameter k	Relative restoration error ρ_{rel}	Estimated discrepancy $\epsilon^2(f,g)$	Estimated object energy E^2
Landweber method	Discrepancy principle	$k = 19$	82.31%	1.99	289.29
	Discrepancy principle (Defrise)	$k = 8$	83.09%	2.62	257.66
	Prescribed energy method	$k = 24$	82.37%	1.85	300.12
Preconditioned Landweber method (Sanz)	Discrepancy principle	$k = 7$	82.49%	1.95	285.94
	Discrepancy principle (Defrise)	$k = 6$	82.44%	2.08	274.52
	Prescribed energy method	$k = 9$	82.75%	1.78	302.85

Table 6.10: Regularization results for object space intensity reconstruction from a *noisy* integral image containing an *opaque cube*, using Landweber and Preconditioned Landweber (Sanz) methods and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

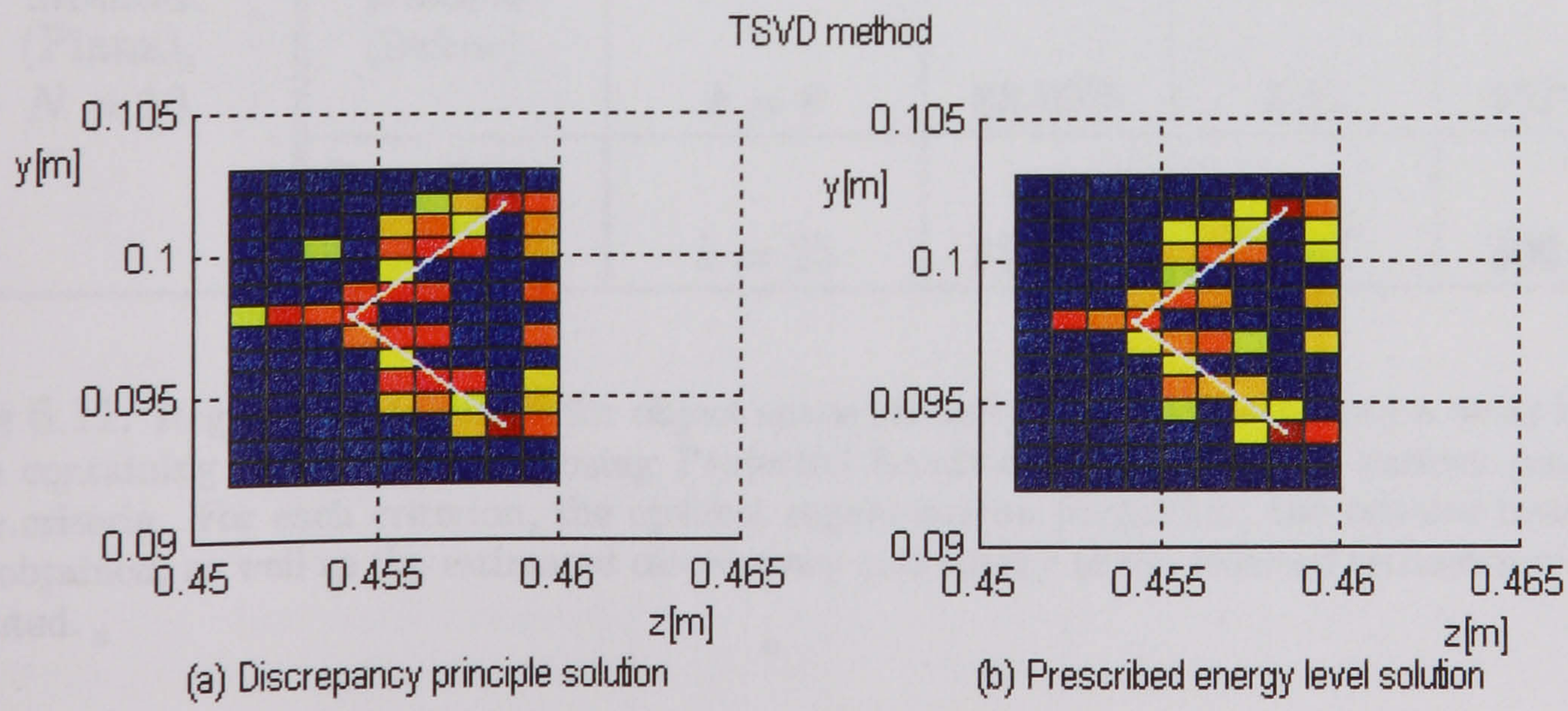


Figure 6.19: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by TSVD regularization and associated parameter choice criteria from a *noisy* integral image.

Method	Stopping rule	Optimal regularization parameter k	Relative restoration error ρ_{rel}	Estimated discrepancy $\epsilon^2(f,g)$	Estimated object energy E^2
Projected Landweber method	Discrepancy principle	$k = 23$	80.74%	2.01	295.24
	Discrepancy principle (Defrise)	$k = 9$	82.00%	2.65	260.32
	Prescribed energy method	$k = 26$	80.76%	1.94	301.40
Preconditioned Projected Landweber method (Sanz)	Discrepancy principle	$k = 8$	81.20%	2.01	292.44
	Discrepancy principle (Defrise)	$k = 7$	81.19%	2.11	283.60
	Prescribed energy method	$k = 10$	81.40%	1.88	306.77
Preconditioned Projected Landweber method (Piana), $N = 10$	Discrepancy principle	$k = 19$	81.76%	2.01	288.39
	Discrepancy principle (Defrise)	$k = 8$	83.09%	2.62	257.00
	Prescribed energy method	$k = 25$	81.46%	1.87	300.09

Table 6.11: Regularization results for object space intensity reconstruction from a *noisy* integral image containing an *opaque cube*, using Projected Landweber methods and various parameter choice criteria. For each criterion, the optimal regularization parameter, the relative restoration error obtained, as well as the estimated discrepancy and energy of the resulted reconstructions are presented.

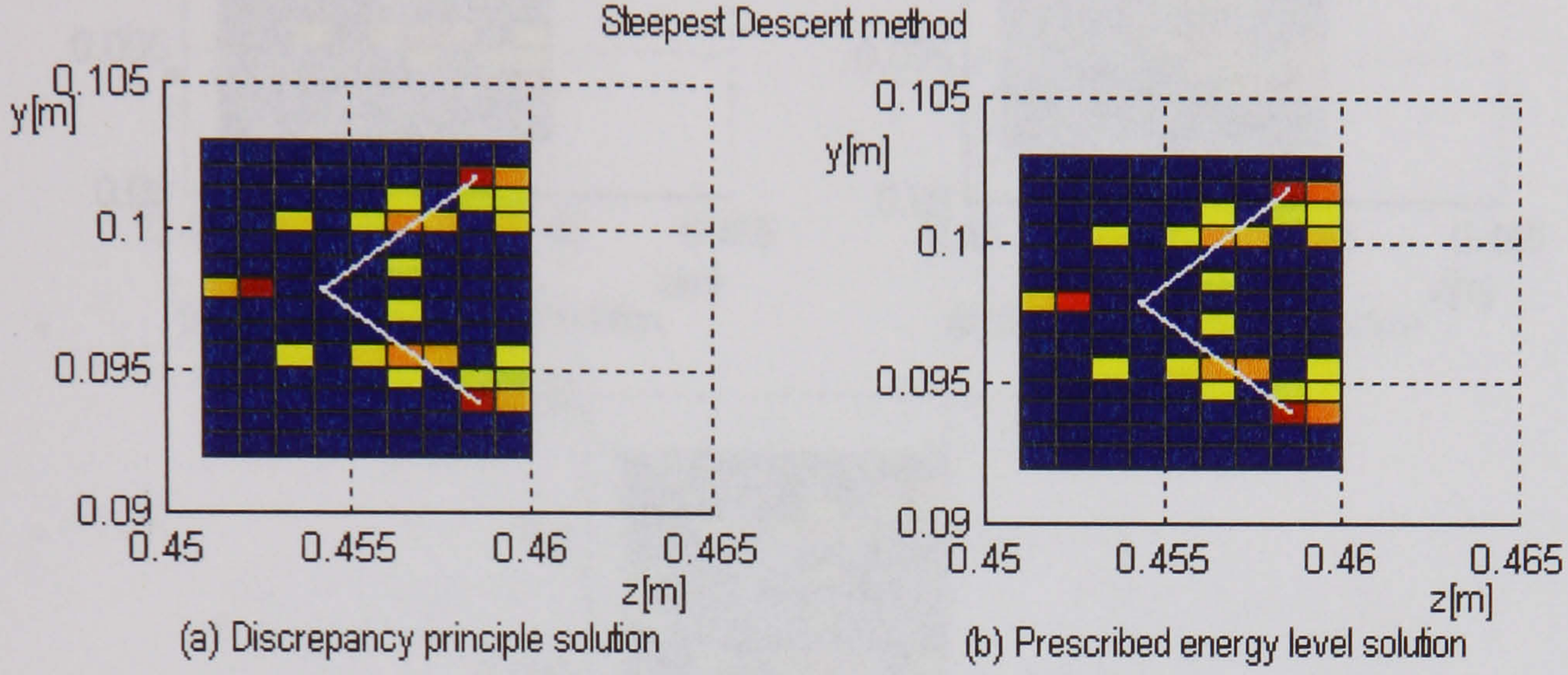


Figure 6.20: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Steepest Descent regularization and associated stopping rules from a *noisy* integral image.

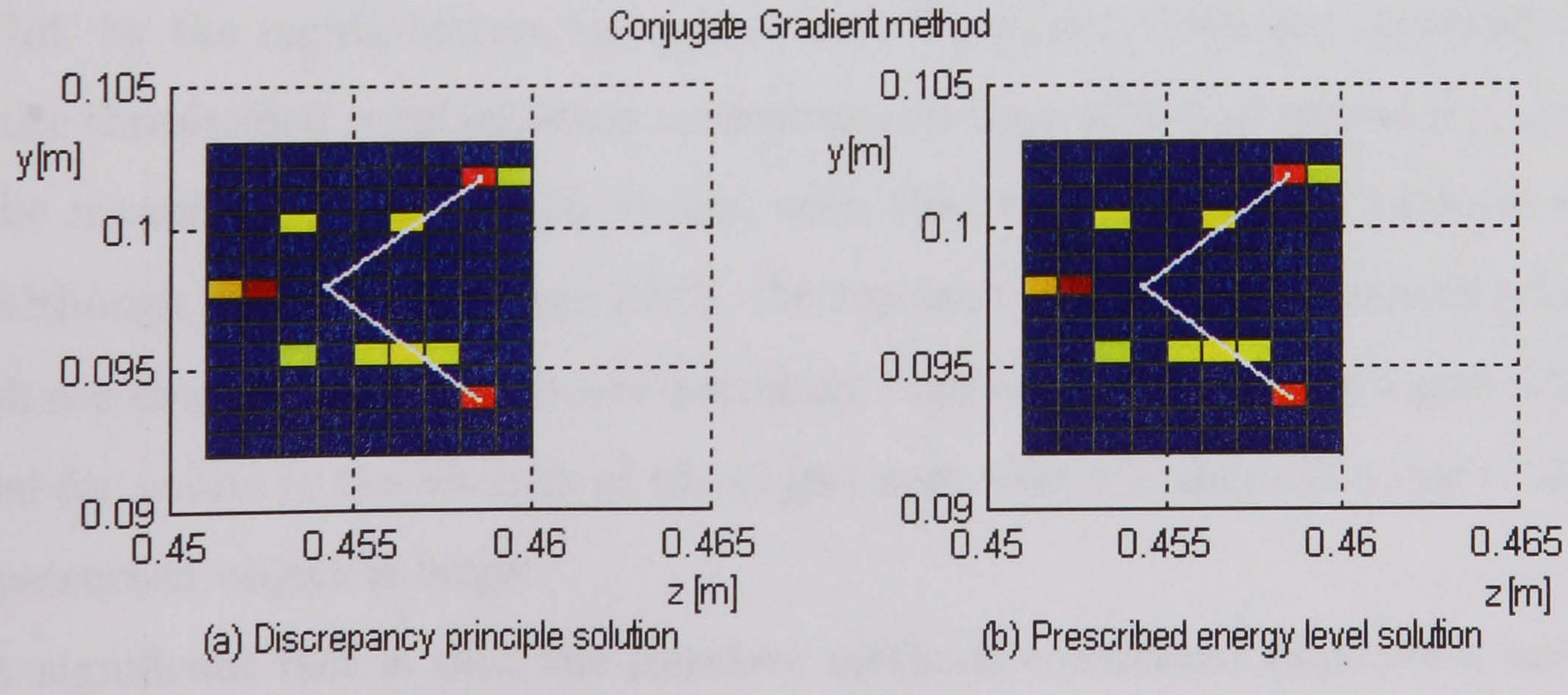


Figure 6.21: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Conjugate Gradient regularization and associated stopping rules from a *noisy* integral image.

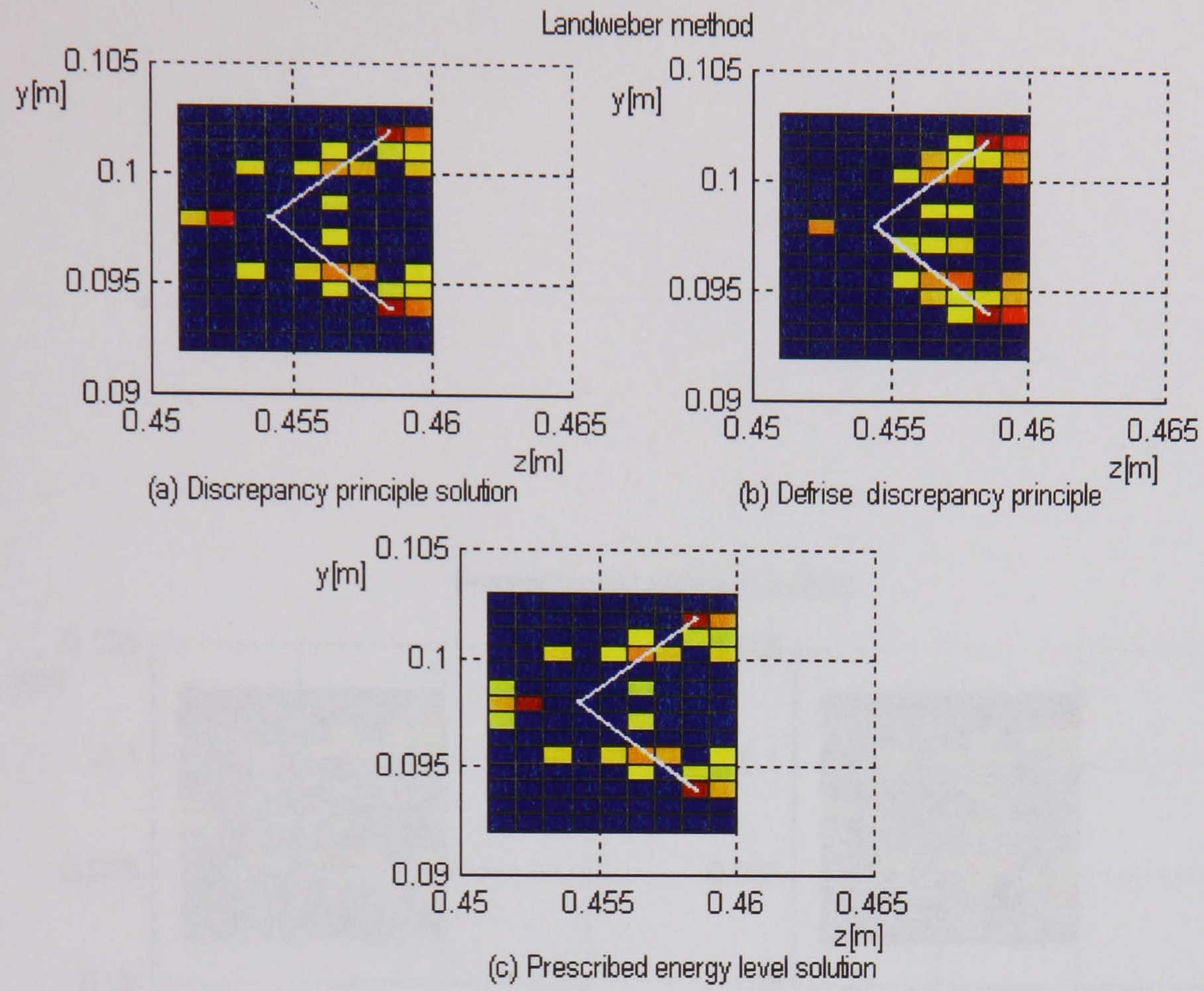


Figure 6.22: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Landweber regularization and associated stopping rules from a *noisy* integral image.

sity (dashed line), and the thresholded normalised version of the object obtained by regularization (continuous line). From the depth extraction point of view, the comparison is relevant if it reveals how many of the expected high intensity points are 'hit' by the regularization reconstruction. Therefore, both the expected object and the thresholded regularization reconstruction have been normalised (i.e. divided by the maximum value in each vector, such that the values range between 0 and 1). Although the match is never 100%, the regularization methods provide solutions which are close to reality. The non-matching high-intensity values are generally obtained for points in the vicinity of the object such that the shape is correct, but the reconstructed object is larger.

A significant fact is that the iterative methods considered required a very low number of iterations to achieve the solution.

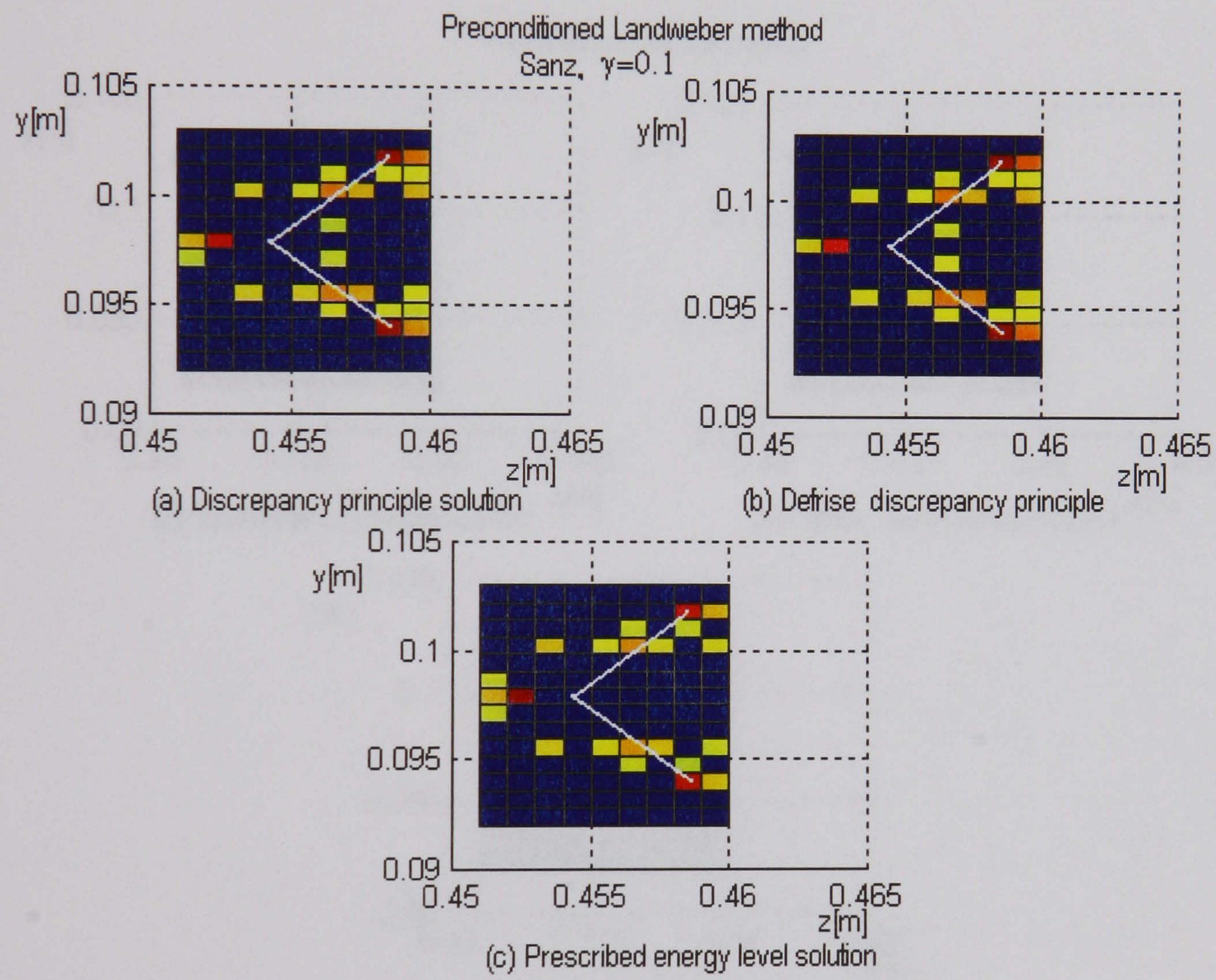


Figure 6.23: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Preconditioned Landweber regularization (Sanz, $\gamma = 0.1$) and associated stopping rules from a *noisy* integral image.

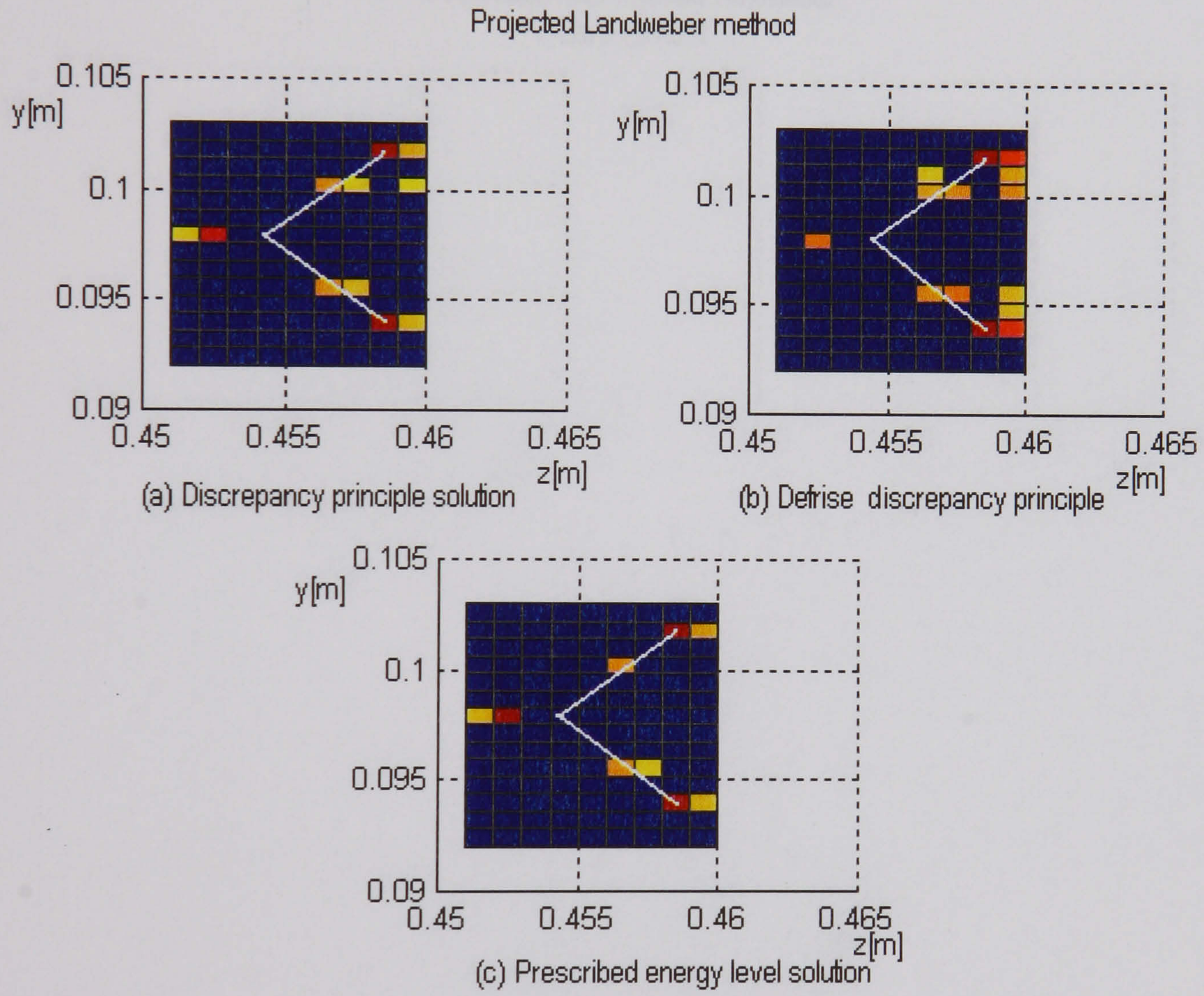


Figure 6.24: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Projected Landweber regularization and associated stopping rules from a *noisy* integral image.

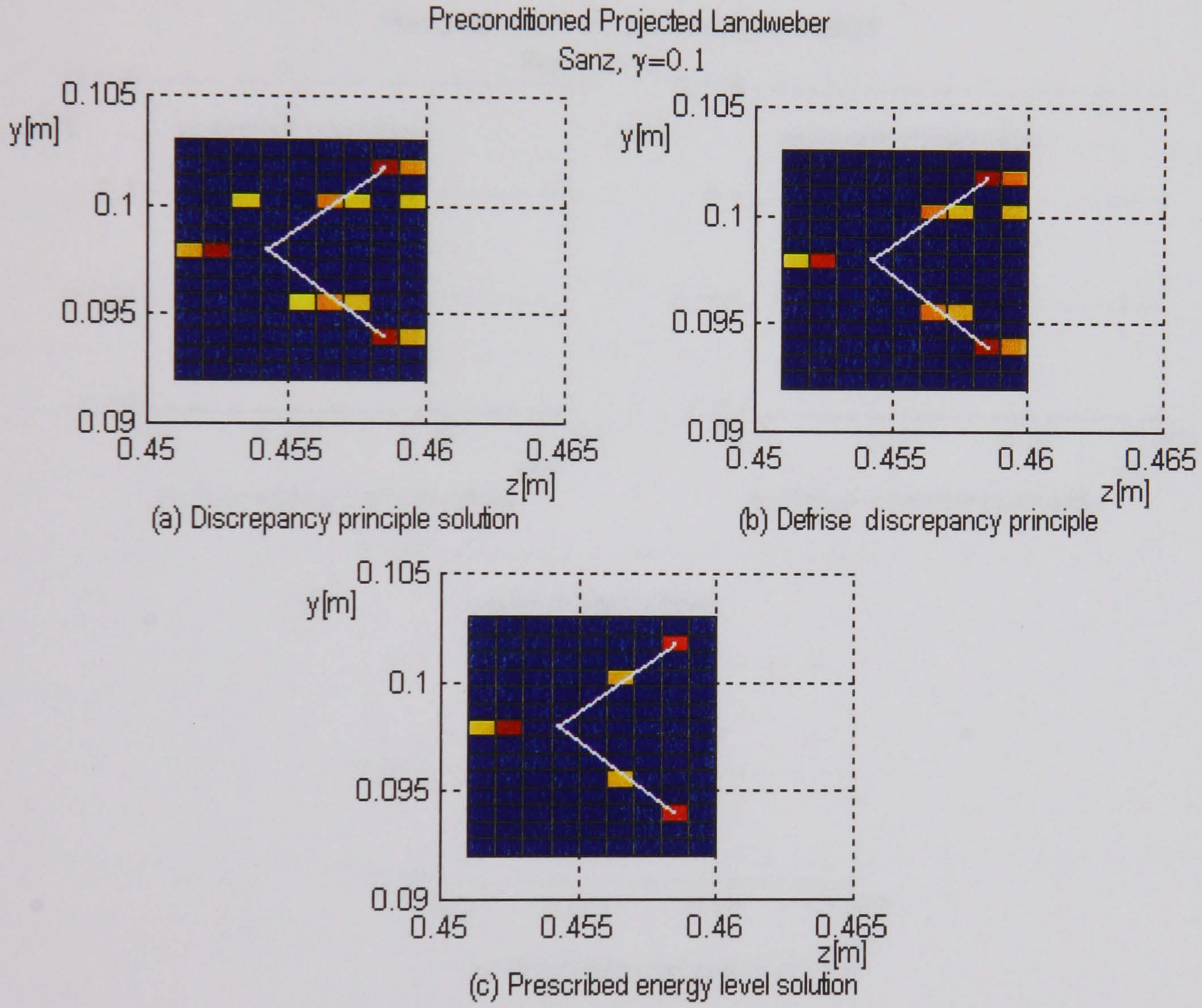


Figure 6.25: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Preconditioned Projected Landweber regularization (Sanz, $\gamma = 0.1$) and associated stopping rules from a *noisy* integral image.

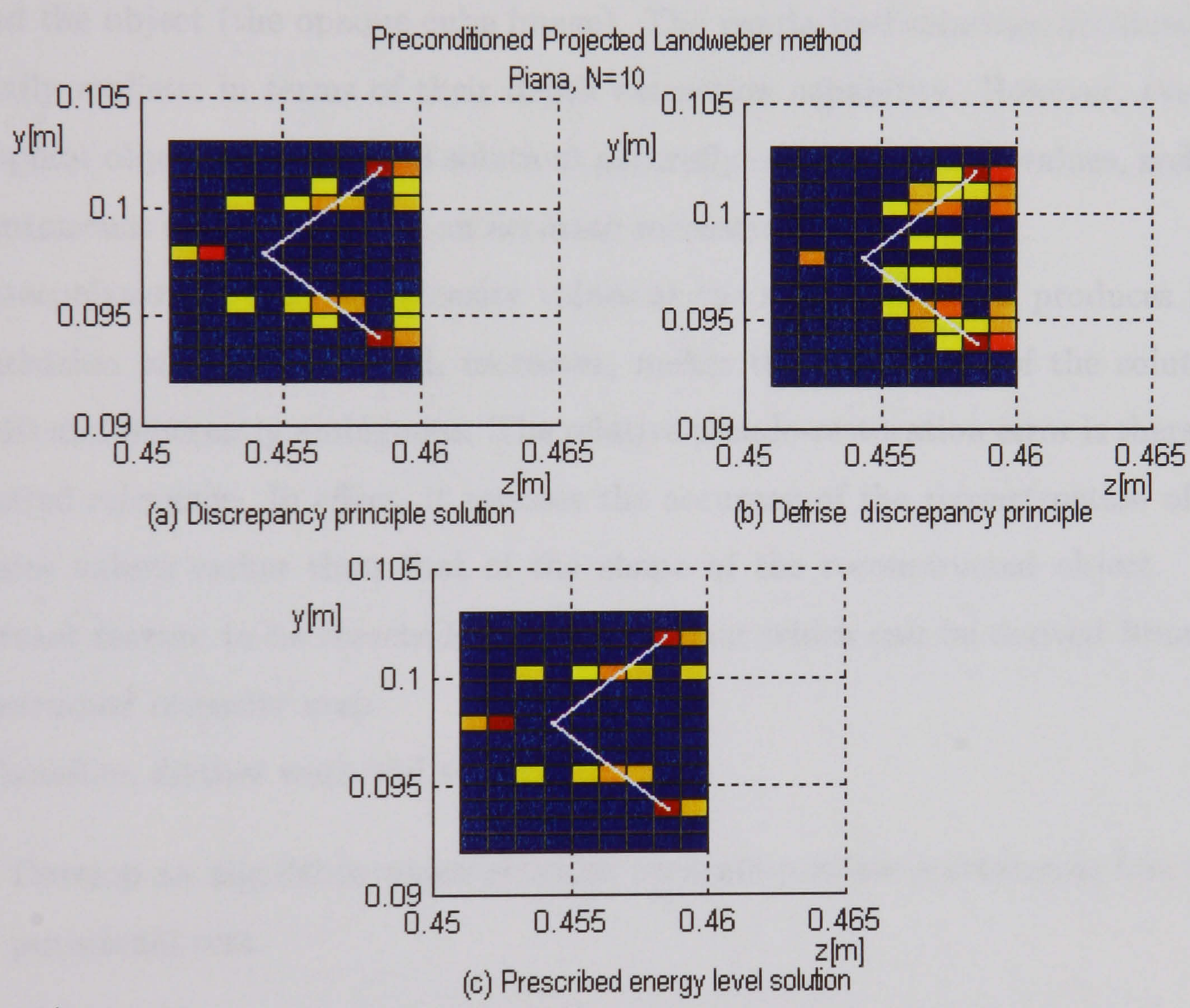


Figure 6.26: Thresholded reconstructions of the visible faces of an *opaque cube* obtained by Preconditioned Landweber regularization (Sanz, $\gamma = 0.1$) and associated stopping rules from a *noisy* integral image.

6.4 Conclusions

The section has presented the application of existing regularization methods to the ill-posed inverse problem of object space reconstruction, and, subsequently, depth extraction from a set of computer generated integral images.

Simulation results have proven that the methods provide very accurate results when some of the sampling points are at or very near the actual object points (the nine luminous point image), but much poorer results when the sampling grid is sparse around the object (the opaque cube image). The regularized solutions obtained are generally realistic in terms of their depth extraction capability. However, even in the 9-point object case, the final solutions generally contain negative values, and the computational cost required by an accurate reconstruction is high.

Interpolation of the real intensity values at the sampling points produces high perturbation of the results and, moreover, makes the assessment of the solutions difficult and inherently ambiguous. The relative pseudo-restoration error is therefore of limited relevance. In effect, it assesses the accuracy of the reconstruction of the intensity values rather than that of the shape of the reconstructed object. The important feature to be searched for is the contour which can be derived from the reconstructed intensity map.

Therefore, further work will seek to:

- Develop an algorithm which provides accurate positive solutions at low computational cost.
- Propose a method of increasing the resolution of the reconstruction.
- Derive a contour, i.e. a depth map, from the reconstructed intensity map.

These issues are the subject of the work reported in the next Chapter.

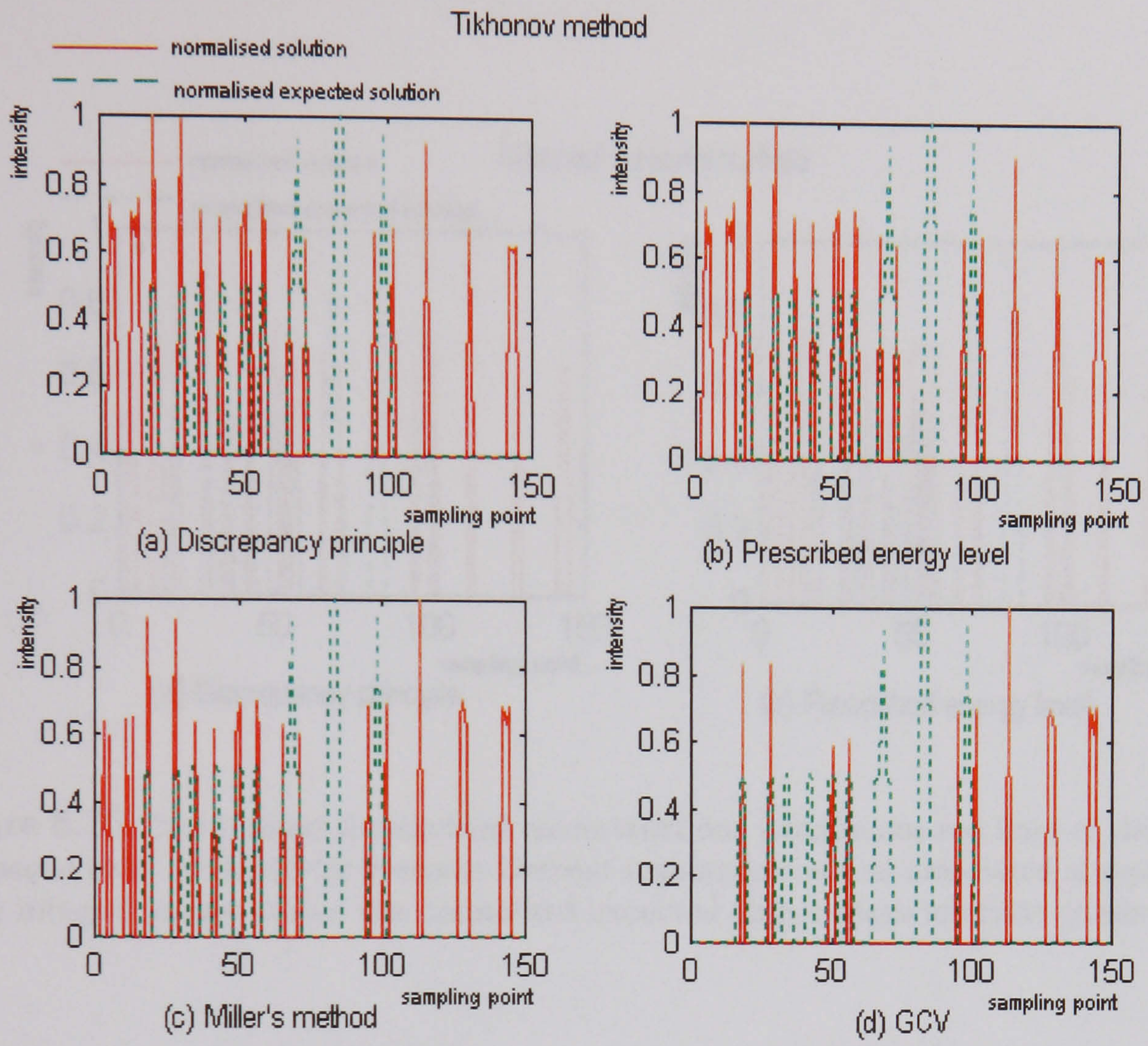


Figure 6.27: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Tikhonov regularization and associated parameter choice criteria from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

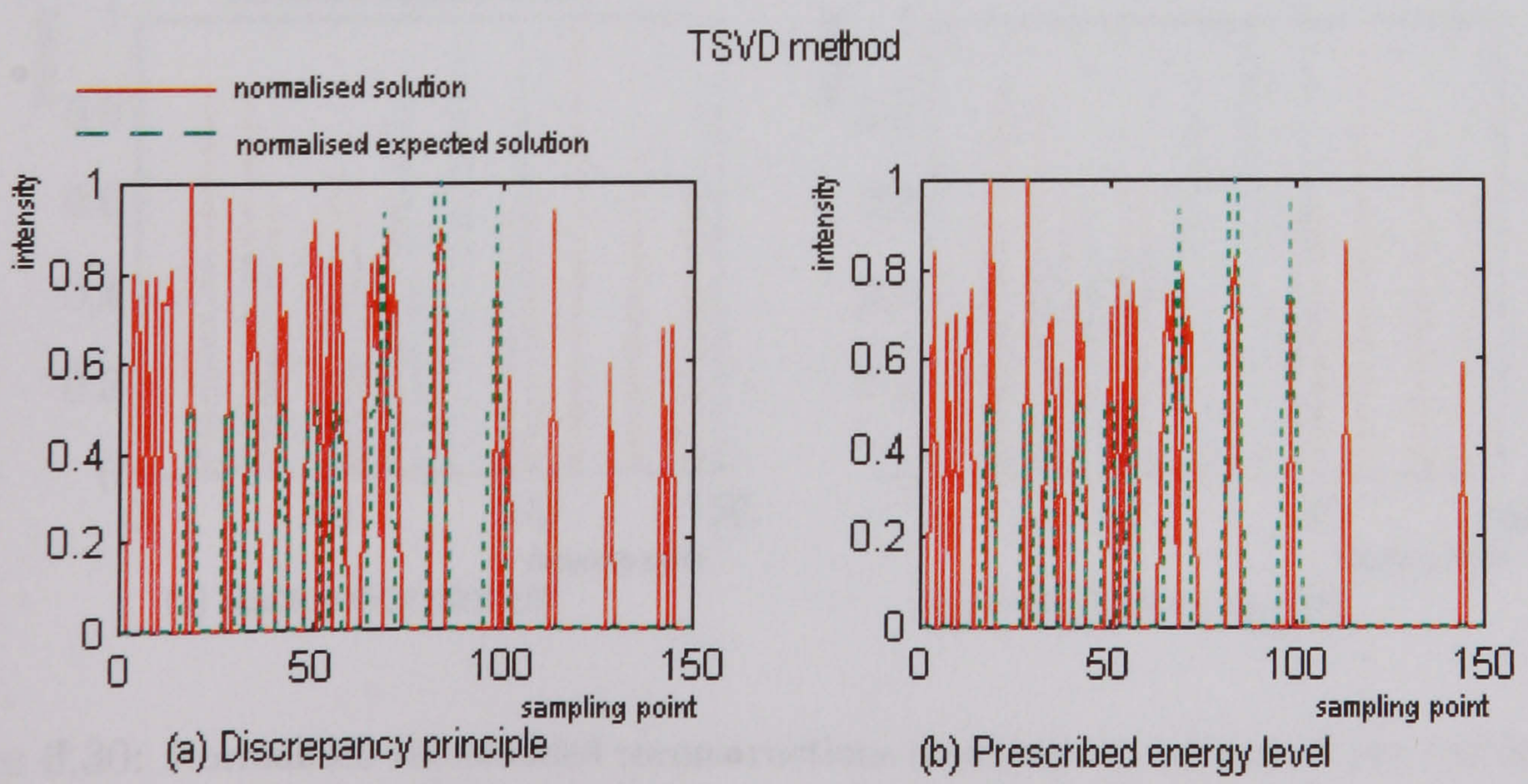


Figure 6.28: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by TSVD regularization and associated parameter choice criteria from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

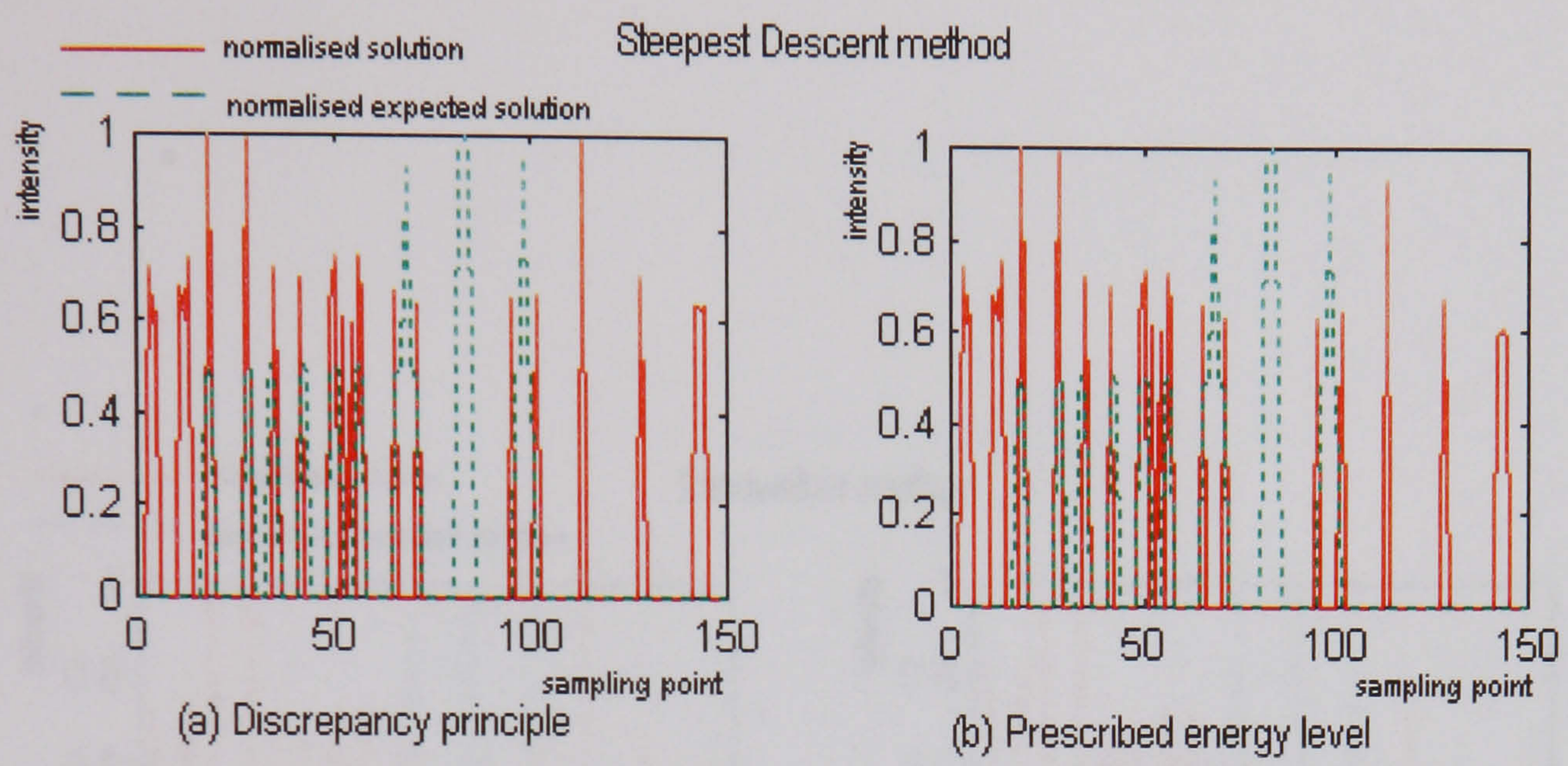


Figure 6.29: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Steepest Descent regularization and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

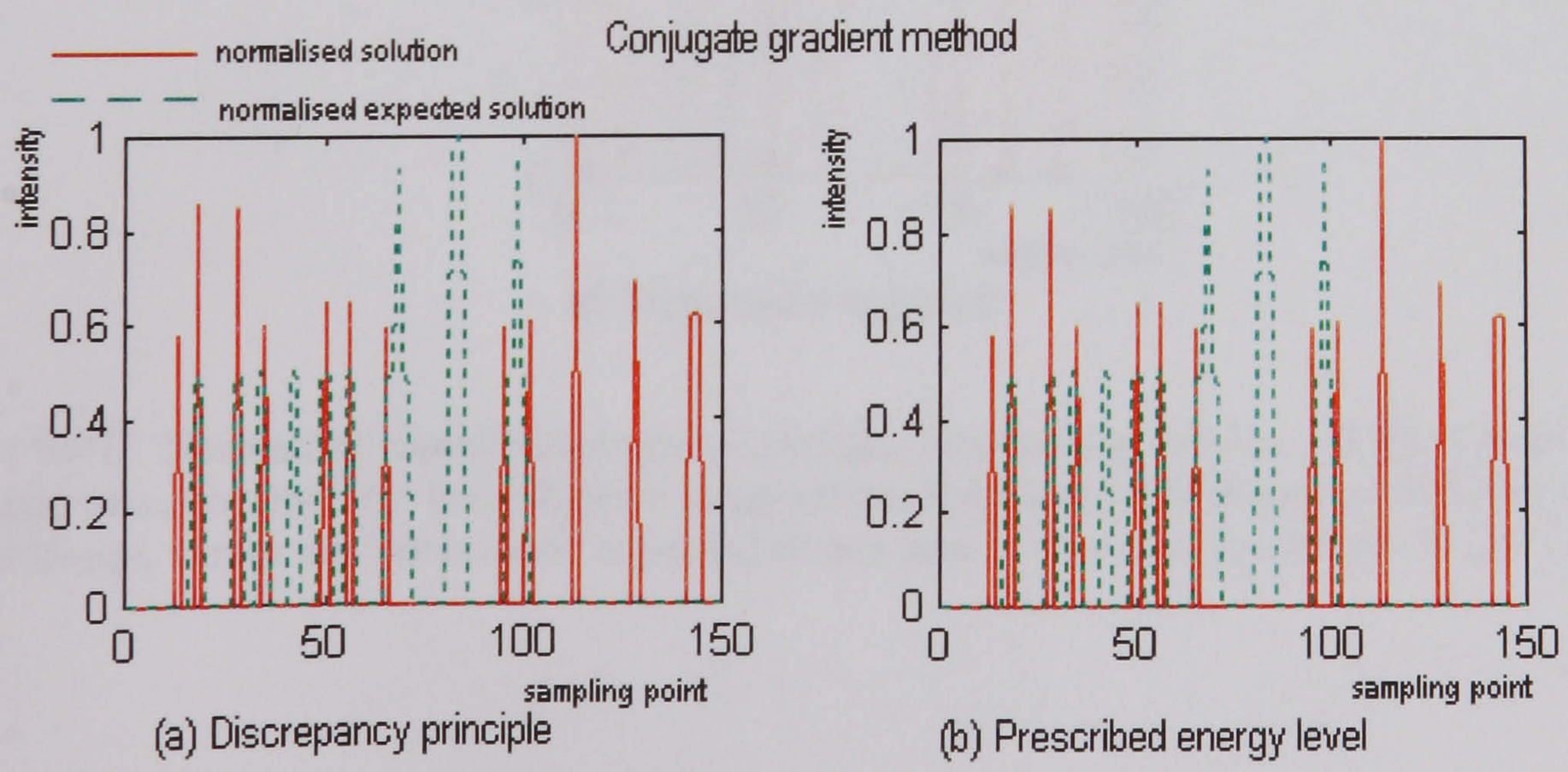


Figure 6.30: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Conjugate Gradient regularization and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

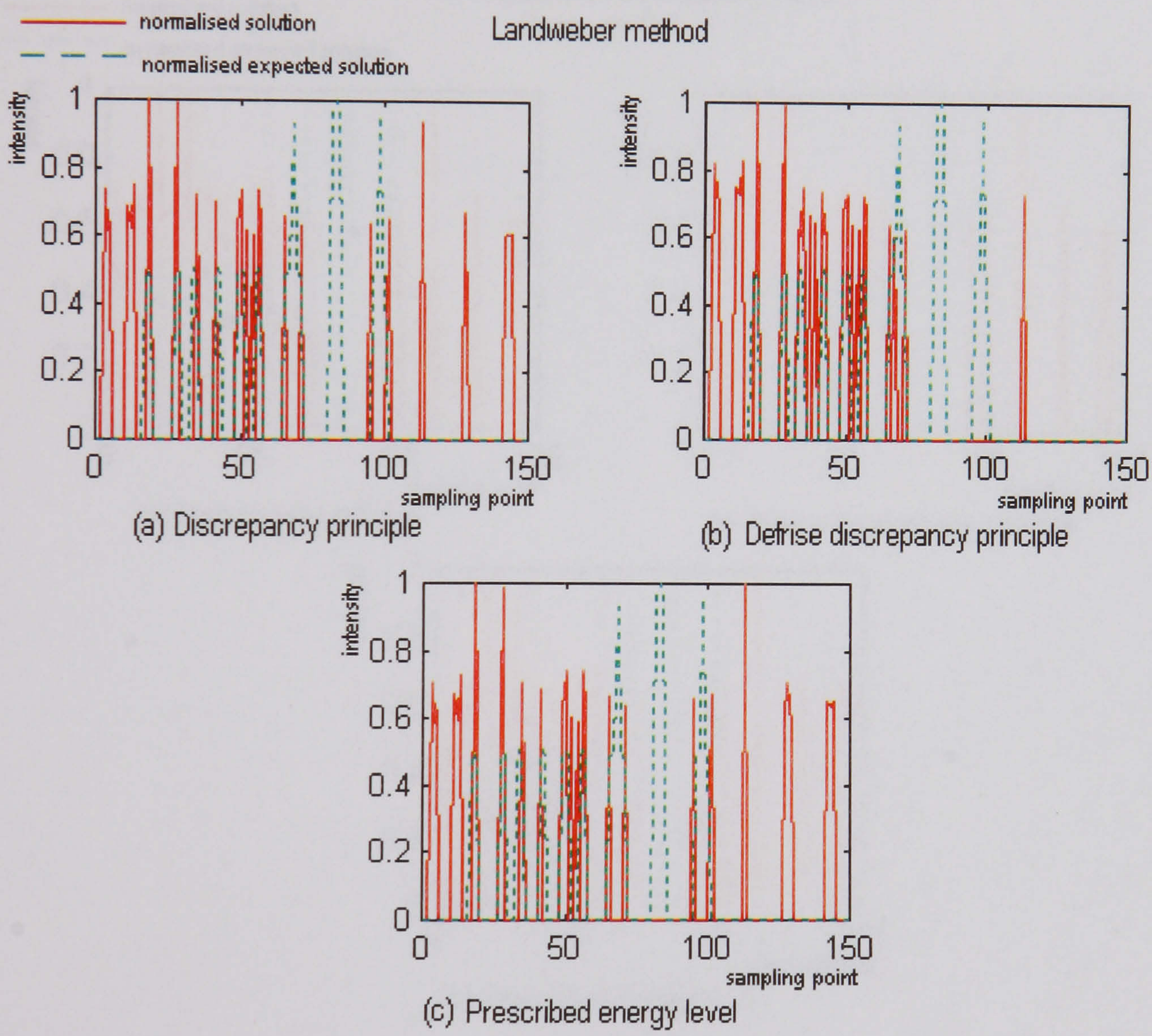


Figure 6.31: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Landweber regularization and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

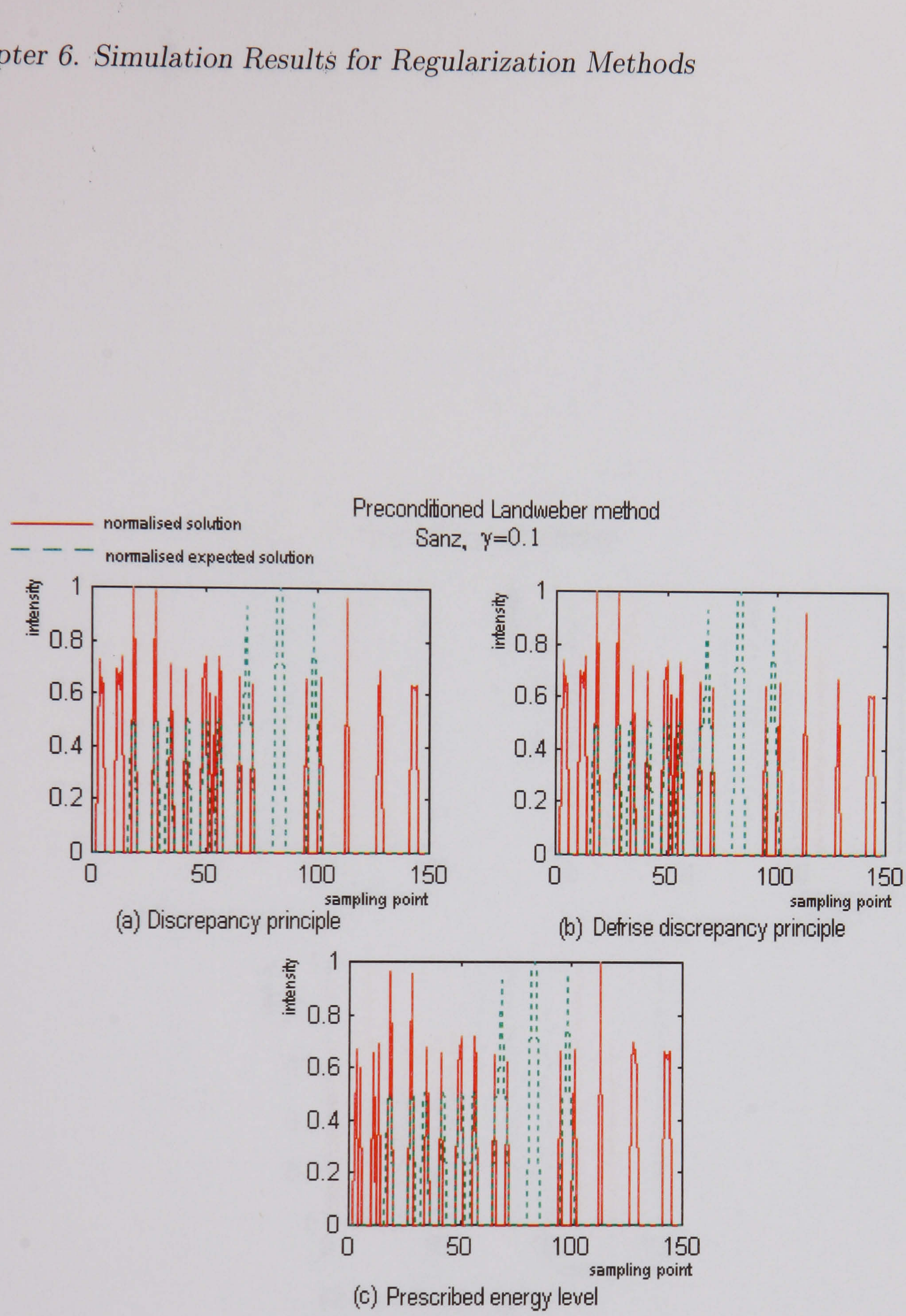


Figure 6.32: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Preconditioned Landweber regularization (Sanz, $\gamma = 0.1$) and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

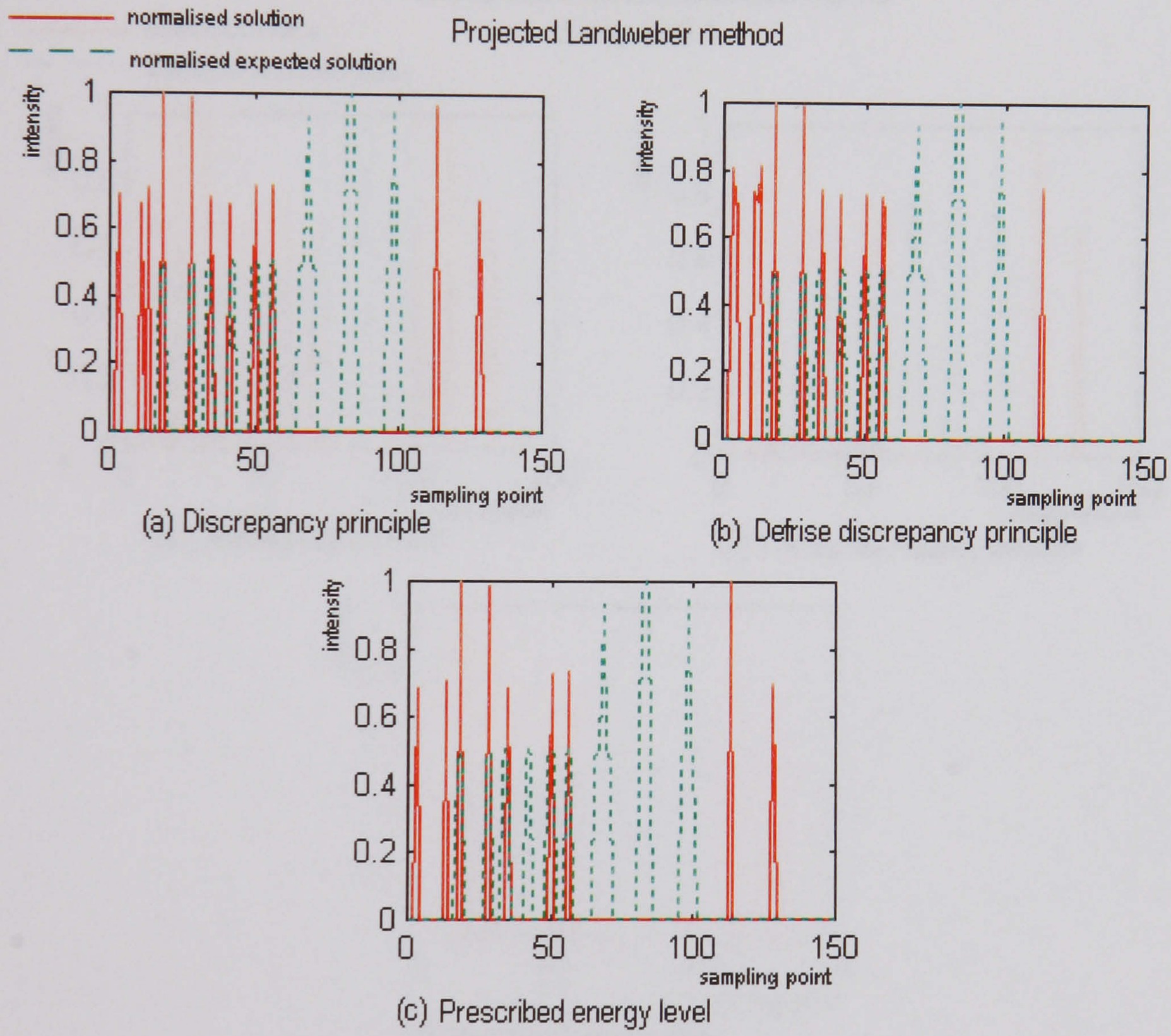


Figure 6.33: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Projected Landweber regularization and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

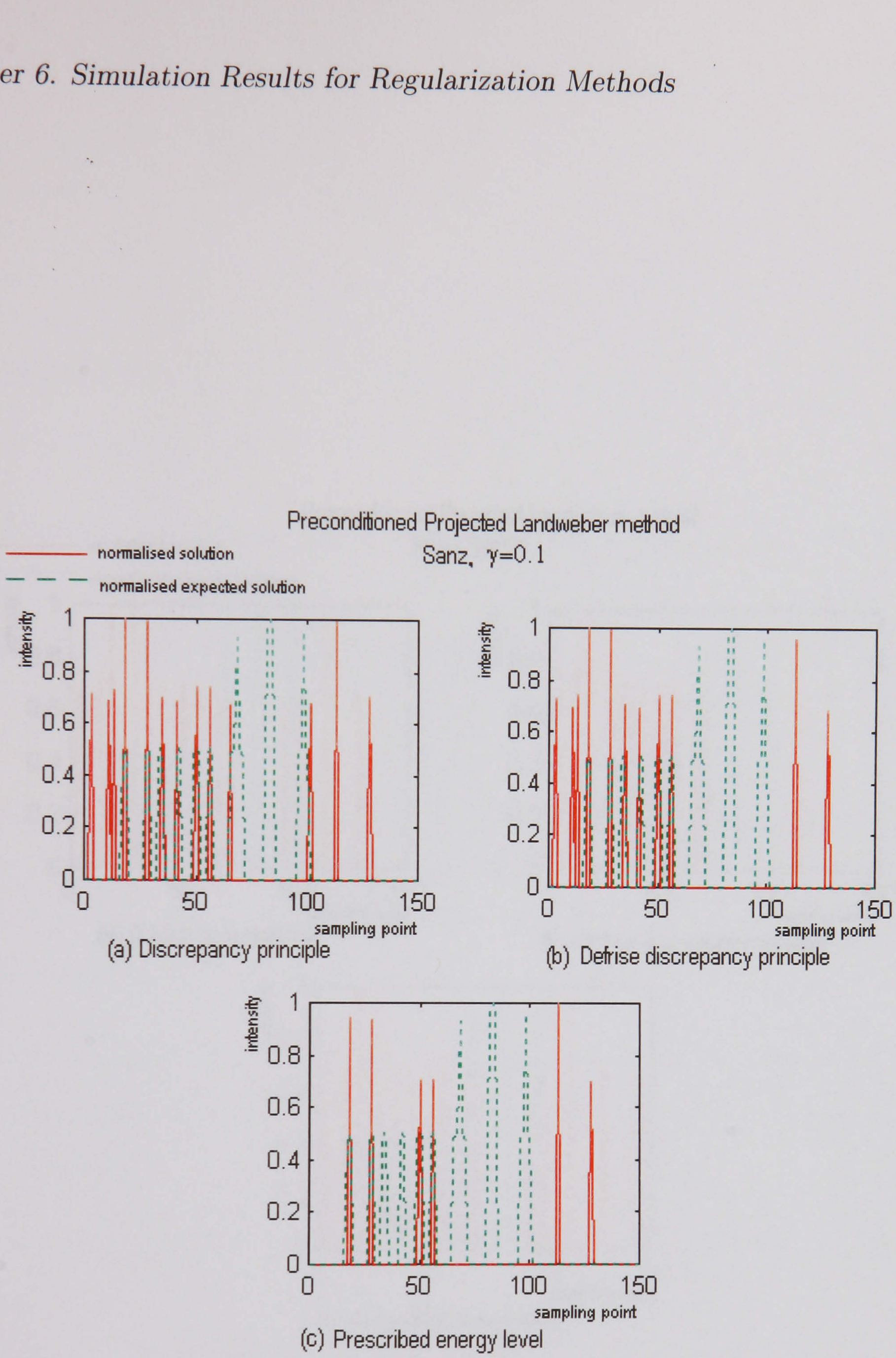


Figure 6.34: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Preconditioned Projected Landweber regularization (Sanz, $\gamma = 0.1$) and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

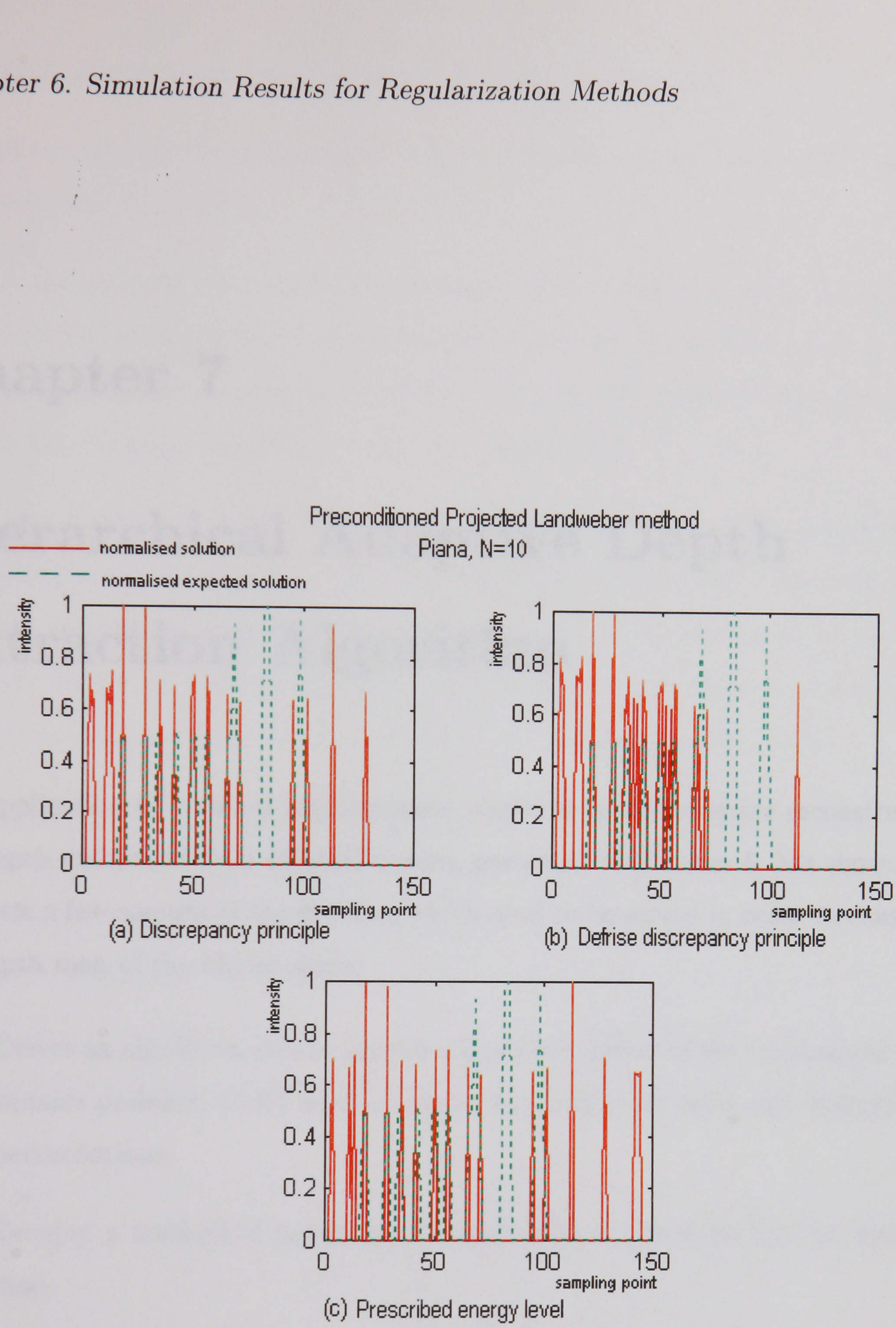


Figure 6.35: Normalised thresholded reconstructions (continuous red line) of the visible faces of an *opaque cube*, obtained by Preconditioned Projected Landweber regularization (Piana, $N = 10$) and associated stopping rules from a *noisy* integral image, versus the normalised expected object space intensity (dashed green line).

Chapter 7

Hierarchical Adaptive Depth Extraction Algorithm

The application of classical regularization methods to object space reconstruction and depth extraction from integral images, presented in Chapter 6, has shown that there are a few aspects of the problem which need to be solved in order to construct the depth map of the object space:

1. Derive an algorithm able to provide a positive solution of the constrained least-squares problem, which is at the same time robust to noise and interpolation perturbations.

2. Develop a method of increasing the resolution of the reconstructed intensity map.

3. Derive a contour, i.e. a depth map, from the intensity map.

The present chapter addresses these problems and proposes a novel algorithm for object space reconstruction and depth extraction from the planar (2D) recorded data set of a 3D-integral image ([58], [59]). This algorithm has three components, each of them responding to one of the stated issues:

1. The space data extraction, formulated in Chapter 5 as an inverse problem, is tackled by imposing additional conditions to the sought solution, and an adaptive constrained 3D-reconstruction (AC3DR) algorithm based on the use

of sigmoid functions is derived. It is very robust to noise and provides positive solutions ([58], [59]).

2. A hierarchical multiresolution strategy (HMS) which employs the adaptive constrained algorithm to obtain higher resolution and higher accuracy intensity maps of the object space is described. The output of the hierarchical algorithm is an *intensity map* of the object space ([58], [59]).

3. The *depth map* of the object space is then extracted from the intensity map using a modified Durbin-Willshaw elastic net. The method, based on a neural network technique, has the salient advantage of being able to overcome the resolution limitations of the reconstructions and to correct inherent reconstruction errors due to interpolation at the sampling points ([59]).

The algorithm has a wide application spectrum which surpasses the boundaries of integral imaging. It can be equally well applied to holography or multi-view systems with an appropriately defined imaging operator. AC3DR can be used for any inverse problem whose sought solutions are positive and bounded. At the same time, the hierarchical multiresolution strategy and the modified Durbin-Willshaw scheme can be used independently in conjunction with any of the existing regularization methods ([60]).

Finally, illustrative simulation results performed on a comprehensive set of test integral images, containing various shapes and scene compositions, are given.

7.1 Adaptive Constrained 3D-Reconstruction Algorithm

The same as for the classical regularization methods, additional constraints coming from a priori knowledge about the object space have to be imposed in order to get a *regularized solution* $\hat{f}$ which corresponds to the physical reality ([6], [88]). More precisely, the solution $\hat{f}$ has to be such that $A\,\hat{f}$ is as close as possible to the image g, i.e.:

$$||A\,\hat{f} - g|| = minimum\,. \tag{7.1}$$

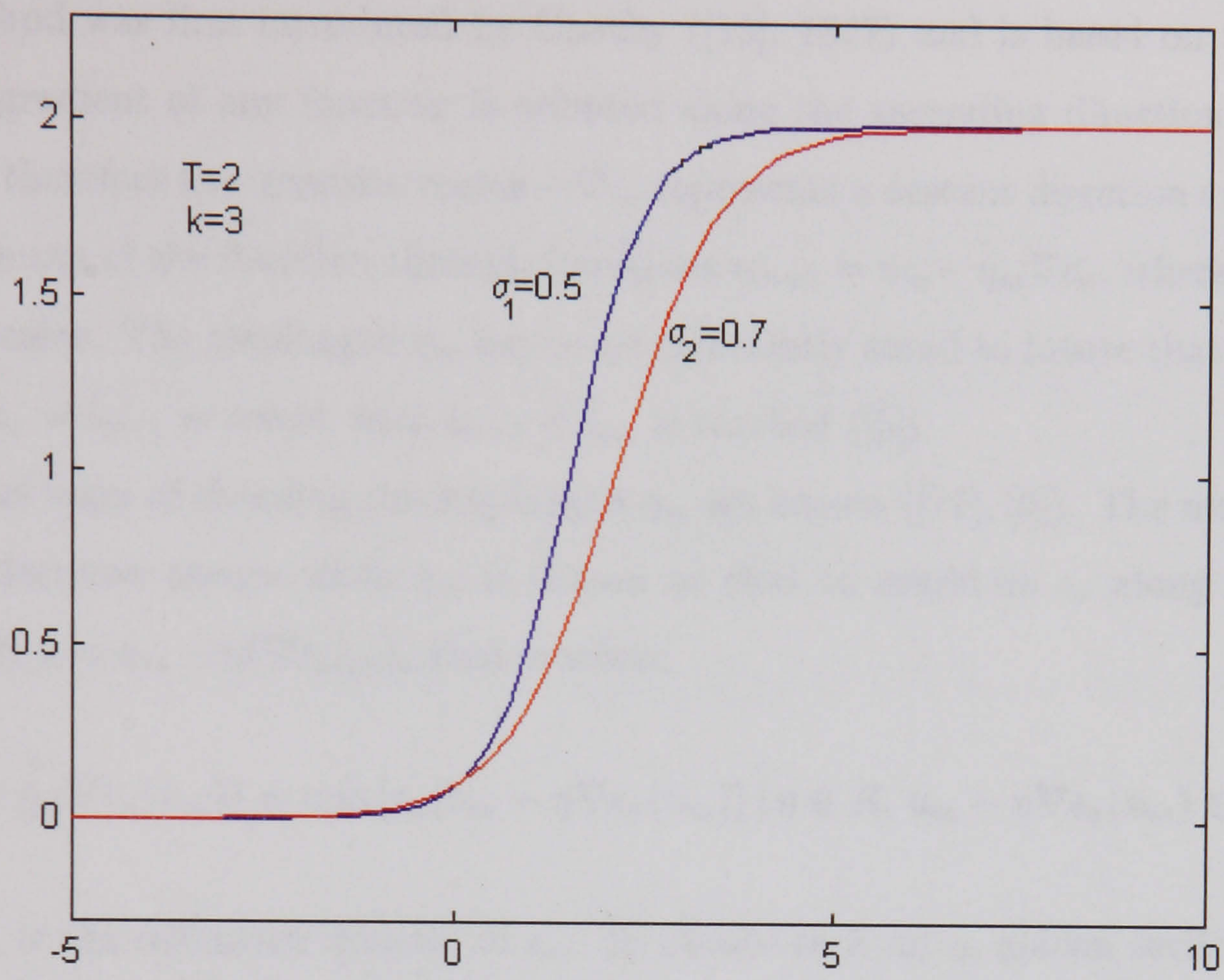


Figure 7.1: Two sigmoid functions.

Also, a condition of positivity and boundedness is necessary, as there are neither negative nor infinite intensity objects:

$$\hat{f} \in [0, T]^n \,. \tag{7.2}$$

$\hat{f}$ satisfying (7.2) can be written as the thresholded value of arbitrary data u: $\hat{f} = f(u)$, where f is a thresholding function. The projection of the interval $[0, T]$ onto R is not differentiable, thus a sigmoid approximation of it, $f_\sigma : R^n \to [0, T]^n$, has been preferred:

$$f_\sigma(u_i) = \frac{T}{1 + exp\left(-\left(\frac{u_i}{\sigma} - k\right)\right)}\,, \; i = 1 \ldots n\,, \tag{7.3}$$

where n is the number of components of the vector $\hat{f}$, i.e. the number of sampling points chosen in the object space. σ is an arbitrary number which controls the slope of the thresholding curve and k is a translation term (Figure 7.1).

To sum up, the sought vector $\hat{f}$ has the form $\hat{f} = f_\sigma(u)$ such that u minimizes the discrepancy functional $\epsilon_\sigma(.) = ||A\, f_\sigma(.) - g||^2$ ([58], [59]).

The gradient descent method ([77], [5]) is used to minimise the discrepancy ϵ_σ. This method was first introduced by Cauchy ([15], 1847) and is based on the idea that the gradient of any function is oriented along the ascending direction of that function, therefore the opposite vector $-\nabla\epsilon_\sigma$ represents a descent direction and leads to a minimum of the function through iterations $u_{m+1} = u_m - \eta_m \nabla\epsilon_\sigma$, where η_m is a positive scalar. The steplength η_m has to be sufficiently small to insure that a lower level set $\epsilon_\sigma = l_{m+1} = const$, with $l_{m+1} < l_m$, is reached ([5]).

Various ways of choosing the steplength η_m are known ([77], [5]). The maximum possible decrease occurs when η_m is chosen so that to minimise ϵ_σ along the ray $\{u \,|\, \eta \in R,\, u = u_m - \eta(\nabla\epsilon_\sigma)_m\}$, that is when:

$$\epsilon_\sigma(u_m - \eta_m \nabla\epsilon_\sigma(u_m)) = \min\{\epsilon_\sigma(u_m - \eta\nabla\epsilon_\sigma(u_m)) \,|\, \eta \in R,\, u_m - \eta\nabla\epsilon_\sigma(u_m) \in D_\epsilon\}\,,$$

where D_ϵ is the definition domain of ϵ_σ. To obtain such an η, golden section algorithm or Brent's algorithm ([5]) can be used.

Another method of choosing the steplength η is provided by the Curry-Altman principle ([77]). It states that η_m is a root of the derivative of the function $\phi(\eta) = \epsilon_\sigma(u_m - \eta\nabla\epsilon_\sigma(u_m))$, that is η_m is the solution of the one-dimensional equation:

$$\epsilon'_\sigma(u_m - \eta\nabla\epsilon_\sigma(u_m))\nabla\epsilon_\sigma(u_m) = 0\,.$$

The gradient method with an optimal value of the steplength η_m has a linear convergence rate, which means that the distance to the minimum point converges to zero as fast as a geometrical progression of sub-unit ratio ([5]).

The gradient of ϵ_σ with respect to σ is ([58], [59]):

$$\nabla\epsilon_\sigma = (A\, f_\sigma(u) - g)^T\, A \begin{bmatrix} f'_\sigma(u_1) & 0 & \dots & 0 \\ 0 & f'_\sigma(u_2) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & f'_\sigma(u_n) \end{bmatrix}. \qquad (7.4)$$

The variation of u which leads to a constrained least-squares solution is on the

descent direction of the gradient $\nabla_\sigma \epsilon_\sigma$:

$$\Delta u = -\eta \nabla \epsilon_\sigma \ , \eta > 0 \ . \tag{7.5}$$

Algorithm 7.1 (Adaptive constrained 3D-reconstruction (AC3DR), [58], [59])

The current solution u_m is modified according to the gradient descent rule:

$$u_{m+1} = u_m - \eta_m \nabla \epsilon_\sigma(u_m) \ ,$$

where η_m is chosen using a suitable steplength algorithm ([77]) such that $\epsilon_\sigma(u_{m+1}) \leq \epsilon_\sigma(u_m)$ and:

$$\nabla \epsilon_\sigma(u_m) = (A \ f_\sigma(u_m) - g)^T \ A \begin{bmatrix} f'_\sigma(u_{m1}) & 0 & \dots & 0 \\ 0 & f'_\sigma(u_{m2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & f'_\sigma(u_{mn}) \end{bmatrix}$$

The algorithm starts with an arbitrary initial solution u_0, e.g. $u_0 = 0$, and is iterated until convergence is reached. The final solution taken into account and which gives the reconstructed intensity map of the object space is $\hat{f} = f_\sigma(u_{final})$. Typical values of the parameters used in simulations are: $T \in [1, 3]$, $\sigma \in [T/2, (3T/2)]$.

The convergence can be tested using one of the following criteria ([5]):

1. Test if the first t significant digits of u_m and u_{m+1} are identical ($\delta_u = 10^{-t}$):

$$\|u_{m+1} - u_m\| \leq \delta_u (1 + \|u_{m+1}\|) \ , \tag{7.6}$$

2. Test if the first t significant digits of $\epsilon_\sigma(u_m)$ and $\epsilon_\sigma(u_{m+1})$ are identical ($\delta = 10^{-t}$):

$$\|\epsilon_\sigma(u_{m+1}) - \epsilon_\sigma(u_m)\| \leq \delta (1 + \|\epsilon_\sigma(u_{m+1})\|) \ , \tag{7.7}$$

3. Test if the gradient is sufficiently small:

$$\|\nabla \epsilon_\sigma(u_{m+1})\| \leq \delta \ . \tag{7.8}$$

This criterion needs to be used carefully as it does not indicate necessarily a minimum point, but also a maximum or a saddle point.

Based on existing local convergence theorems([77]), it can be proven that there exits a steplength $\eta_i > 0$ at each iteration such that the algorithm converges to a local solution for any initial value $u_0 \in (-M, k\sigma)^n$, where M is a large fixed real number ([59]).

The solution obtained with the adaptive constrained algorithm is an approximation of the global unique solution I^* whose uniqueness has been stated in Theorem 5.1.

Though it has not been proven yet that this algorithm satisfies the regularization definition in the sense of Tikhonov ([88]), it will be termed as 'regularization method' in the following because it provides a minimal discrepancy solution of the problem.

'1 out of 3 bins' thresholding rule ([59]). Finally, the *regularized solution* given by the AC3DR algorithm is *thresholded automatically* using a 3-equidistant bin histogram. The values contained in the lowest bin, i.e. the 'background', are set to 0 and the rest of the values are kept (Figure 7.2). It differs from the '2 out of 3 bins' rule used for thresholding the solution obtained with classical regularization methods in Chapters 5-6. Numerical simulations have shown that this 3 bin histogram thresholding rule provides reliable solutions for the current type of object reconstruction. Typical histograms with 10 and 3 equidistant bins of an AC3DR regularized solution are presented in Figures 7.2(c)-(d). It can be observed that the thresholding rule adopted is very natural given the structure of the reconstruction. At the same time, Figures 7.2(a)-(b) present the histograms of the scene reconstruction from the same integral image and with the same sampling net using Tikhonov regularization. The layout of the histograms indicates that the background and the edges in the AC3DR regularization are much better differentiated from each other than in the reconstruction obtained with a classical regularization method, where the reconstructed intensity values are uniformly distributed. This is an important advantage of the new algorithm because it facilitates the selection of the contour in the object space.

Another important feature of the method is the very low sensitivity to noise. The reconstructions obtained from pictures with an SNR=30 dB differ in norm with

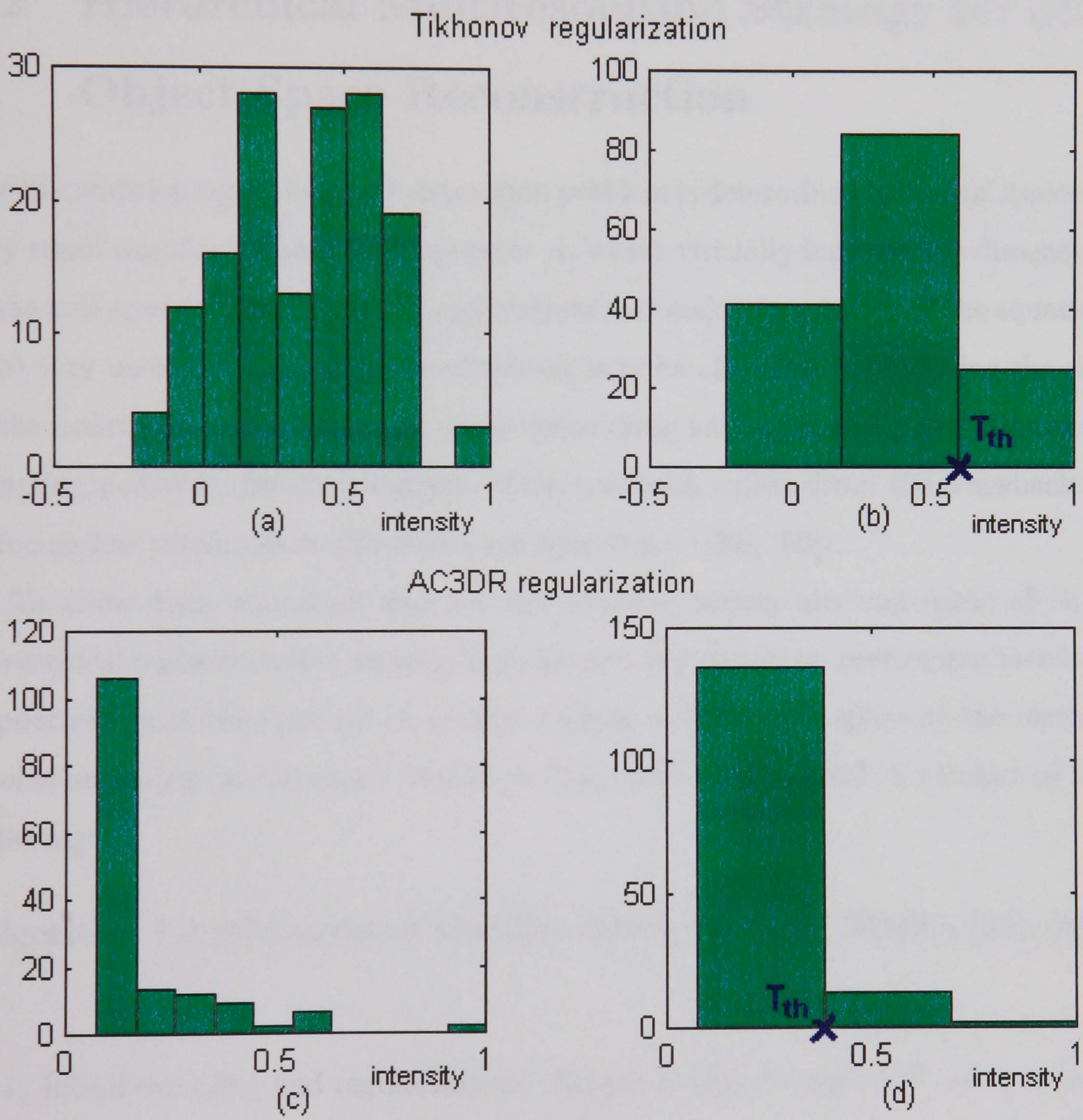


Figure 7.2: 10- and 3- equidistant bin histograms of solutions obtained using Tikhonov regularization ((a) and(b)) and AC3DR regularization ((c) and (d)). T_{th} represents the value chosen as a threshold.

1-2% from the reconstructions obtained from the noiseless pictures, whilst images with an SNR=20 dB give reconstructions that differ from those obtained in the absence of noise with 5-6%. We can therefore conclude that the method is very robust to noise.

7.2 Hierarchical Multiresolution Strategy for 3D-Object Space Reconstruction

The ill-conditioning of the depth extraction problem is determined by the existence of very small singular values of the operator A, which virtually increase the dimension of the null-space of A in numerical applications and make the solution of the equation (5.5) very unstable. Such an ill-conditioning may be alleviated by reducing the size of the matrix corresponding to A, so by considering only a relatively low number of sampling points in the object space. This operation suffers from the drawback of inducing low resolution in the object reconstruction ([59], [58]).

To allow finer resolution and yet not increase unduly the null-space of A, a *hierarchical multiresolution strategy* is proposed. It is based on performing localised regularization of the solution on certain regions of the object space at the current resolution and/or at increased resolution ([59], [58]). This approach consists of the following:

Algorithm 7.2 (Hierarchical Multiresolution Strategy (HMS), [59], [58])

1. Initial sampling and regularisation: *Sample a large domain* $D^{(1)}$ *using an initial grid* Σ_1 *and obtain a reconstruction* $\hat{f}_{D^{(1)}}$ *from the equation* $g = A^{\Sigma_1}_{D^{(1)}} \hat{f}^{\Sigma_1}_{D^{(1)}}$ *using AC3DR algorithm.*

2. Localized regularization: *Select the high intensity regions* D_1,.., D_m *of the object space from the current reconstruction* $\hat{f}^{\Sigma_1}_{D^{(1)}}$ *and consider the rest of the space as determined. Perform local regularization of the reconstruction solution on the union* $D^{(2)} = D_1 \cup \ldots \cup D_m$ *using one of the following alternatives:*

(a) Local regularization at the same resolution: *Consider the restriction of the sampling Σ_1 to the current domain $D^{(2)}$ and obtain a more accurate solution $\hat{f}^{\Sigma_1}_{D^{(2)}}$ from the equation $g = A^{\Sigma_1}_{D^{(2)}} \hat{f}^{\Sigma_1}_{D^{(2)}}$ using the AC3DR algorithm.*

(b) Local regularization at increased resolution: *Resample the current domain $D^{(2)} = D_1 \cup \ldots \cup D_m$ using a finer grid Σ_2 and obtain a finer resolution reconstruction $\hat{f}^{\Sigma_2}_{D^{(2)}}$ from the equation $g = A^{\Sigma_2}_{D^{(2)}} \hat{f}^{\Sigma_2}_{D^{(2)}}$ using the AC3DR algorithm.*

3. Further resolution refining: *If necessary, step 2 can be recursively applied until the object space is reconstructed at the desired resolution, which should not exceed the resolving capability of the optical system.*

Remark. Simulation results have shown that a suitable resolution choice is essential for obtaining reliable regularized solutions. Oversampling a region at a resolution which exceeds the resolution capability of the physical system leads to a more blurred reconstruction of the object, though correctly shaped and positioned ([59]). An illustrative example is presented in section 7.4.

The output of the hierarchical application of the adaptive constrained regularization is an *intensity map* of the object space. This map allows the segmentation of a multiple object scene into separate components. The segmentation is facilitated by the hierarchical structure of the algorithm ([59], [58]).

Note. The hierarchical multiresolution strategy can be used independently of the AC3DR algorithm, in conjunction with any of the existing regularization methods ([60]) to increase the resolution and accuracy of the reconstruction.

7.3 Depth Map Generation

An important problem which needs to be addressed is to generate the depth map, i.e. the contours in the object space, from the existing intensity map. The automatic thresholding of the intensity map obtained with the hierarchical adaptive 3D-reconstruction algorithm provides 'edges' and 'boundaries' of the object contained in the imaged scene, and an intelligent procedure to thin and link them is

needed. The solution chosen to perform these tasks in the case of depth extraction is based on a neural network technique, namely the use of a Durbin-Willshaw elastic net, modified to increase the 'attraction' or 'gravitation' of the higher intensity reconstructed points ([26]).

Traditional edge thinning and linking are generally associated with high resolution images and confined to the pixelated structure of the image ([4]). For depth mapping, a solution based on a neural network technique has been preferred instead because it has the outstanding advantage of being able to overcome the resolution limitations of the reconstructions and to correct inherent reconstruction errors due to interpolation at the sampling points ([59]). As already emphasised, interpolation poses an important problem at object space reconstruction, irrespective of the reconstruction method used. However, important information is retained by the intensity values of the reconstructed points. The distribution of this information *as a whole* is used by the modified Durbin-Willshaw scheme, which is a global method, to produce the final depth map *among* the sampling points. It needs to be stressed that the depth map output by the method is *not* attached to the sampling grid used for reconstruction. From a mathematical point of view, the points on the Durbin-Willshaw contour are floating points in a *continuous object space*, not points (or voxels) of the sampling grid. In this way, both resolution and interpolation problems are compensated ([59]).

In the following, a theoretical layout of the Durbin-Willshaw scheme for the travelling circus problem is given, and the way it is adapted and applied to depth map generation is presented.

7.3.1 Travelling Circus Problem and Elastic Net Solution

The Durbin-Willshaw elastic net was originally designed to solve travelling circus problems ([26]):

Given l cities, find the optimal positions of the m stops of a circus such that the efficiency of the tour is maximum.

In the Durbin-Willshaw solution of the problem ([26],[46]), an initial path is gradually elongated non-uniformly according to an iterative procedure until it eventually passes near to all the cities to define an optimal tour. Each point defining the

path moves under the influence of two types of constraints:

- the point is constrained to move towards those cities which are nearest;
- the point is pulled towards its neighbours on the path such that the total path length is minimised.

In this way, each city becomes associated with a certain section on the path. The tightness of the association is determined by how the influence contributed by a city depends on the distance, and the nature of this dependence changes as the algorithm progresses. Initially, all the cities have the same influence on each stop on the path. Progressively, longer distance becomes less favoured, and each city gradually becomes more specific to the stops on the path closest to it.

Suppose that p_ξ, $\xi = 1, \ldots, l$, are the positions of the cities and that the stop path is defined by an ordered set of points $w = (w_1, \ldots, w_m)$. If the two constraints mentioned are taken into account, the following energy functional of the entire net w can be defined ([26]):

$$E(w) = -\varrho^2 \sum_{\xi} \log[\sum_{j} \exp(-|p_\xi - w_j|^2/(2\varrho^2))] + \frac{\alpha}{2} \sum_{j} |w_{j+1} - w_j|^2 \ . \qquad (7.9)$$

In the equation, ϱ expresses the effective attraction radius of each city p_ξ, and α is a factor which controls the mutual attraction between two consecutive stops on the path.

The first term of the sum (7.9) represents the Gaussian attraction which allows each city to stimulate the move of a stop towards it. The extent of this attraction is controlled by the radius ϱ. The second term expresses the tension between neighbouring stops on the path and helps minimise the contour length, assuring that the consecutive w_js are reasonably close to each other.

The optimal path w is one with minimum energy E. The convergence towards a minimum energy path is realised using a gradient descent method. Thus, having initialised the net, it deforms to link all the cities, producing contours which are usually within a few percent of known minimal tours ([35]).

The updating rule is the following:

$$w_i = w_i - \eta \nabla w_i \,, \tag{7.10}$$

with $\eta > 0$ and ([26], [46]):

$$\nabla w_i = \sum_{\xi} \Lambda^{\xi}(i)(p_\xi - w_i) + \alpha(w_{i+1} - 2w_i + w_{i-1}) \,, \tag{7.11}$$

where $\Lambda^{\xi}(i)$ is defined as:

$$\Lambda^{\xi}(i) = \frac{\exp(-|p_\xi - w_j|^2/(2\varrho^2))}{\sum_j \exp(-|p_\xi - w_j|^2/(2\varrho^2))} \,. \tag{7.12}$$

7.3.2 Weighted Durbin-Willshaw Scheme for Depth Map Generation from an Intensity Map

The final purpose of the present work is to draw the *depth map* of the object space, i.e. the contour of the scene, from the intensity map provided by the hierarchical adaptive constrained 3D-reconstruction algorithm. Since it has been shown that the edge point linking is isomorphic to the travelling circus problem ([35]), this algorithm makes a good candidate for contour and depth extraction from thresholded intensity maps of scene reconstructions from integral images. As the intensity maps produced contain only the vertices and a few points along the edges, this technique is able to extract the original object. Boundary points are read in from the intensity map and become the cities in the travelling circus problem. The technique aims to produce the shortest contour among those points, which typically coincides with the actual outline of the object.

To draw the depth map of the 3D-scene, the following procedure has been adopted ([59]):

1. The 3D-intensity map of the scene is derived as a set of voxels using the hierarchical adaptive 3D-reconstruction algorithm and the automatic thresholding rule.

2. For each horizontal plane of voxels $x = x_0$, an optimal curve, defined by an

ordered set of points $w = (w_1, \ldots, w_m)$, which fits the high intensity reconstructed points is sought using a Durbin-Willshaw scheme.

The intensity map may contain voxels (or points) having various values, and obviously a differentiation between points according to their intensity needs to be done. In the classical form of the Durbin-Willshaw algorithm, all the 'cities' have equal importance and influence to the net. Therefore, a modification of the scheme is necessary to encompass the intensity differences between two voxels ('cities'). Each voxel (or 'city') is assigned a rate or weight $r(p_\xi)$, which is a number proportional to the square of the intensity of the point p_ξ. The energy of the whole net becomes in this case ([59]):

$$\tilde{E}(w) = -\varrho^2 \sum_\xi \log[r(p_\xi) \sum_j \exp(-|p_\xi - w_j|^2/(2\varrho^2))] + \frac{\alpha}{2} \sum_j |w_{j+1} - w_j|^2 \ , \quad (7.13)$$

where p_ξ , $\xi = 1, \ldots, l$, are the high intensity points obtained at thresholding, ϱ determines the effective attraction radius of each point p_ξ, and α assures that the consecutive w_j are reasonably close to each other. The gradient changes to the following expression ([59]):

$$\nabla w_i = \sum_\xi r(p_\xi) \Lambda^\xi(i)(p_\xi - w_i) + \alpha(w_{i+1} - 2w_i + w_{i-1}) \ . \quad (7.14)$$

The weighted Durbin-Willshaw algorithm for minimising $\tilde{E}$ by gradient descent and obtaining the contour of the object in a plane $x = x_0$ is:

Algorithm 7.3 (Weighted Durbin–Willshaw scheme, [59])

1. - *Consider the high intensity points* p_ξ , $\xi = 1, \ldots, l$ *and their associated intensities* i_ξ , $\xi = 1, \ldots, l$. *Compute the rating coefficients:* $r(p_\xi) = (i_\xi / \min_\xi(i_\xi))^2$, $\xi = 1, \ldots, l$.
   - *Consider an arbitrary initial solution* $w^{(0)} = (w_1^{(0)}, \ldots, w_m^{(0)})$.

2. *Update the current solution according to the gradient descent rule:*

$$w^{(k+1)} = w^{(k)} - \eta \nabla w^{(k)} \ ,$$

where $\eta > 0$ *and* $\nabla w^{(k)}$ *is given by:*

$$\nabla w_i^{(k)} = \sum_{\xi} r(p_\xi)\Lambda^{\xi}(i)(p_\xi - w_i^{(k)}) + \alpha(w_{i+1}^{(k)} - 2w_i^{(k)} + w_{i-1}^{(k)}) \ ,$$

for any $i = 1, \ldots, m$.

Step 2 is iterated until convergence is reached. Typical values of the parameters used in simulation are: $\varrho = 0.001 \div 0.003$ for the attraction radius of a city, and $\alpha = 1 \div 3$ for the stops' mutual attraction factor.

In the case of multiple object scenes, the contour of each object can be drawn using a separate weighted Durbin–Willshaw structure.

The weighted Durbin-Willshaw algorithm has been found to be extremely robust and stable. It is a global technique and generally links the high intensity points irrespective of their location in the plane. Moreover, it demonstrates tolerance to initial configuration problems. It can be used with intensity maps produced by any of the object space reconstruction methods (hierarchical AC3DR or the regularization methods presented in Chapter 5).

7.4 Simulation results

The hierarchical adaptive constrained 3D-reconstruction algorithm has been tested on a set of computer generated unidirectional integral images, noiseless or noisy, containing various object shapes and depths (Figures 8.2-8.8, Appendix C). The results obtained for images containing a 9 disparate point configuration, a thin face having various inclinations with respect to the camera aperture, a two transparent thin face scene, a transparent cylinder, and an opaque cube are presented in the following. Some of these results have already been reported in [58] and [59]. The noisy versions of the pictures were obtained by corrupting the noiseless pictures with white Gaussian noise of mean 0 and variance equal to the square of 1-5% of the maximum intensity value contained by the noiseless picture. The images thus obtained have a signal to noise ratio around 20-30 dB.

For the simplicity of the presentation, planar sections parallel to the Oyz-plane of the object space reconstruction are shown. The objects considered are all per-

pendicular to the Oyz-plane, so their intersections with $x = x_0$ planes are similar. Hence, recovering one section means recovering the whole object. The colour scaling in the reconstructions is determined by the intensity range obtained for the sampling grid, so the colour is proportional to the intensity of the point. Warm colours correspond to high intensity values, while cold colours represent low intensity values.

EXAMPLE 7.1 *The reconstruction of a 9-point scene*

The computer generated image of a scene containing nine luminous disparate points situated at various (known) depths and lateral positions in space has been processed. Both a noiseless version and a noise contaminated version have been considered. They are shown in Figure 8.2, Appendix C. All nine object points are situated in the same horizontal plane of the object space, which has been sampled with a 150 point grid that contains the exact positions of the nine object points. The AC3DR regularization performed with this sampling on the noisy image, thresholded with the 3 bin histogram rule, provided the reconstruction shown in Figure 7.3, which reproduces exactly the original points. The grid was further compressed to cover a smaller area around the nine points at higher resolution and the result was the same: the nine points were reproduced accurately.

In this particular case where the exact locations of the object points are among the sampling points, it is possible to compare the reconstruction and the original object, and to calculate the relative restoration error. The normalised reconstructions obtained from the noiseless and noisy images of the nine points are shown versus the original object space intensity in Figures 7.4 and 7.5 respectively. The error is quasi-insignificant in both cases: $\rho_{rel} = 4.7\%$ for the reconstruction from the noiseless image, and $\rho_{rel} = 4.46\%$ (so smaller!) for the reconstruction from the noisy image. The fact that the relative error obtained for the noisy picture is lower than that obtained with the noiseless picture indicates that the method is very robust to noise, which is a major feature of the algorithm.

The scene contains disparate points, so the intensity map obtained by hierarchical regularization is also a depth map.

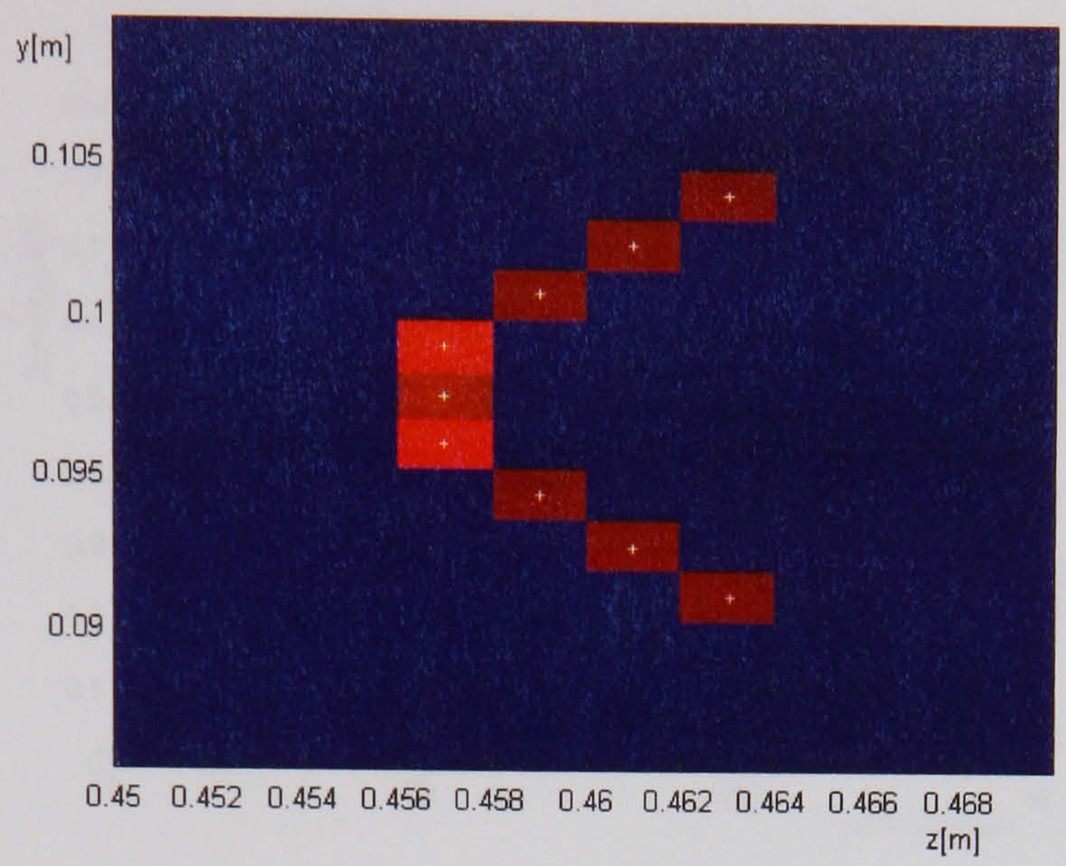


Figure 7.3: AC3DR algorithm reconstruction of a *nine point object* scene from a *noisy* integral image. The true positions of the object points are represented by the white crosses.

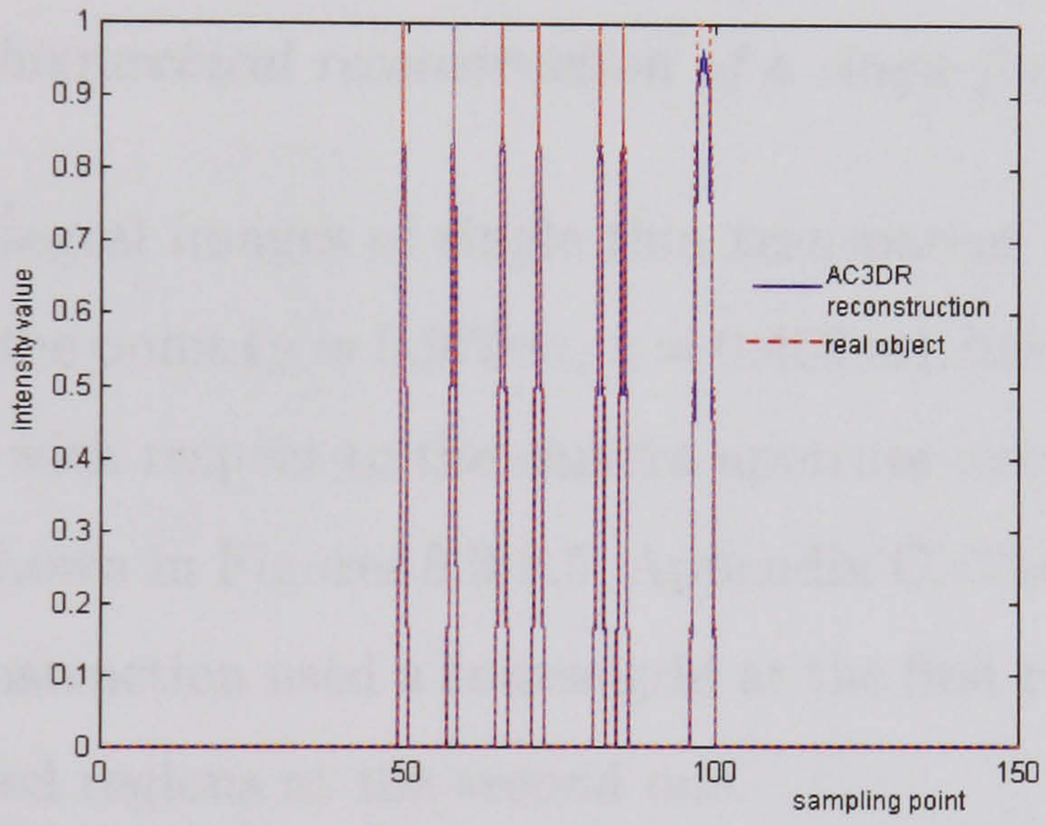


Figure 7.4: Normalised thresholded AC3D reconstruction (continuous blue line) of a *nine point object* from a *noiseless* integral image versus the object space intensity (dashed red line).

This is a 'toy' example which shows that the procedure works very well in the ideal case when the object points are among the sampling points. In the case of continuous objects, like the ones concerned in the following examples, the solution is an interpolation of the real intensity values at the sampling points. This is why the Durbin–Willshaw scheme is necessary to decide on the final reconstructed contour of the object space.

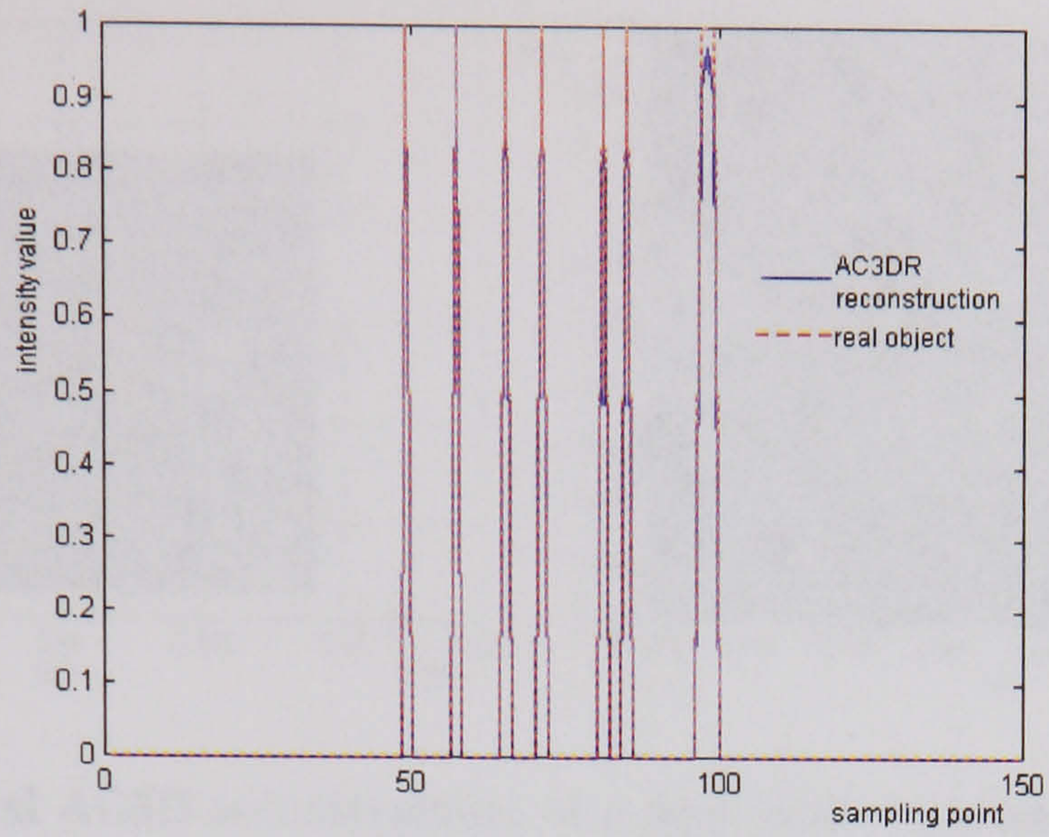


Figure 7.5: Normalised thresholded AC3D reconstruction (continuous blue line) of a *nine point object* from a *noisy* integral image versus the object space intensity (dashed red line).

EXAMPLE 7.2 *The hierarchical reconstruction of a single face*

Unidirectional integral images of single thin transparent faces, 0.4 mm wide, 8 mm long, centred at the point ($y = 0.975\,m$, $z = 0.459\,m$), having various inclination angles (0°, 45°, 90°) with respect to the camera aperture were also considered. The integral images are shown in Figures 8.3-8.5, Appendix C. The hierarchical adaptive constrained 3D-reconstruction used a course grid at the first regularization level and a finer grid on selected regions at the second one.

- *First hierarchical level: initial sampling and regularization.* The reconstructions at low resolution are shown in Figure 7.6(a) for a thin face parallel to the camera aperture, in Figure 7.7(a) for a face inclined at 45° with respect to the camera aperture, and in Figure 7.8(a) for a face perpendicular on the camera aperture. The objects are detected at their true size and correctly positioned in the reconstruction. The face orthogonal to the camera is very well reproduced, while the other two present perturbed shapes. The high intensity region is selected in each reconstruction and used at the next regularization phase.

- *Second hierarchical level: localised regularization with increased resolution (algorithm 7.2.2b).* Further processing is performed on the marked regions around the high intensity regions selected at the first stage using finer sampling grids, and more accurate reconstructions are obtained (Figures 7.6(b), 7.7(b), 7.8(b))

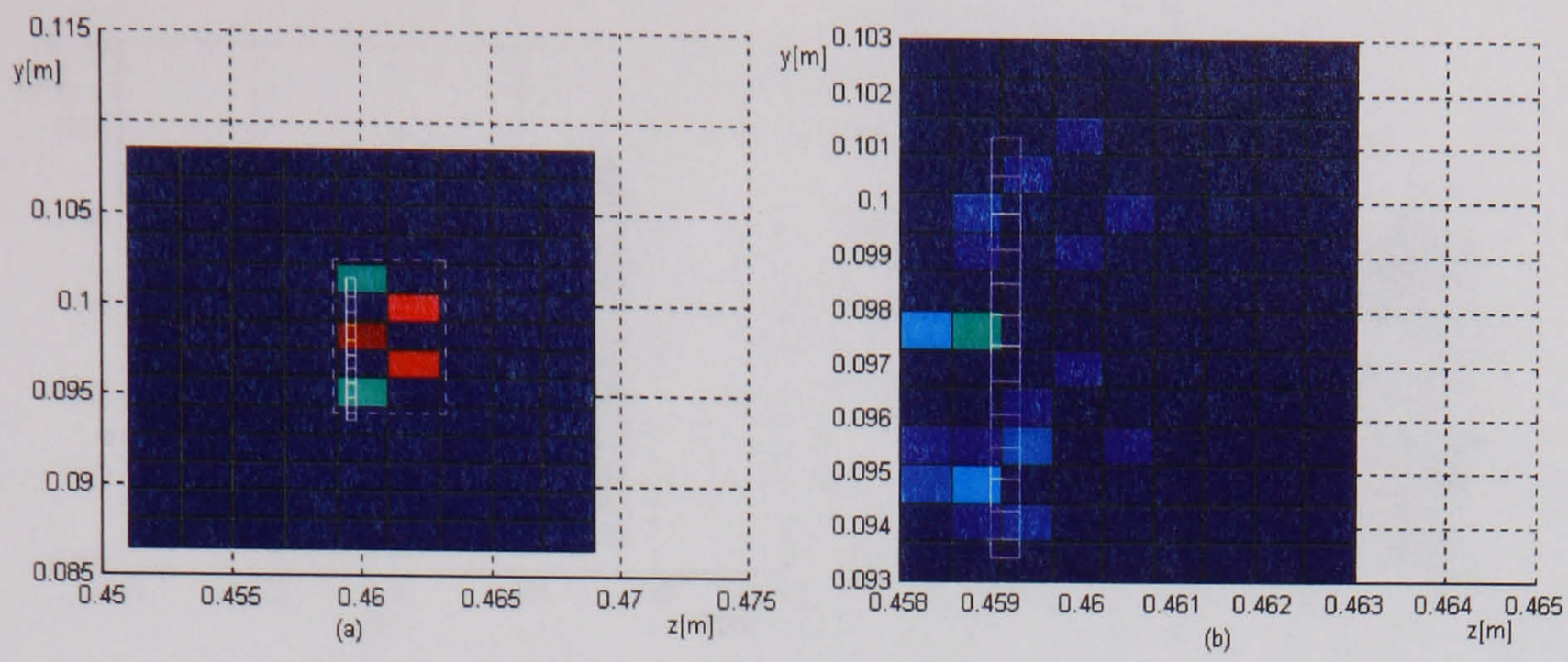


Figure 7.6: Hierarchical AC3D reconstruction of a *face* (plane) parallel to the camera aperture from an integral image: (a) initial regularization and high intensity region selection (dashed line); (b) regularization at increased resolution in the region selected at (a). The original object is figured in continuous line.

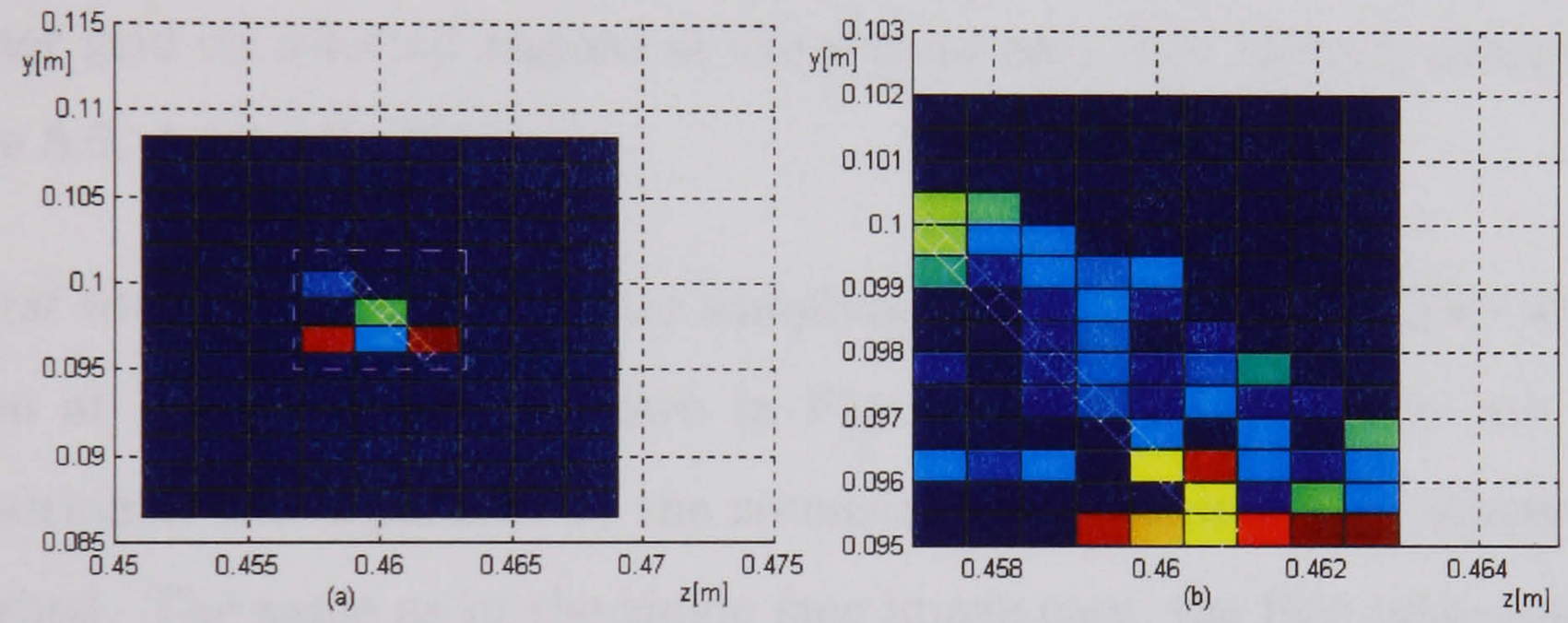


Figure 7.7: Hierarchical AC3D reconstruction of a *face* (plane) which forms a 45 degree angle with the camera aperture: (a) initial regularization and high intensity region selection (dashed line); (b) regularization at increased resolution in the region selected at (a). The original object is figured in continuous line.

using the *multiresolution* approach. Again the face orthogonal to the camera aperture is well reproduced as position, shape, and dimensions. The 45° inclined face has a blurred reconstruction, but which respects the position and dimensions of the true object. The reconstruction of the face parallel to the camera aperture presents lack of definition, but good positioning and lateral dimension.

EXAMPLE 7.3 *The hierarchical reconstruction of a two object scene*

The object scene contains two thin transparent faces of unequal lengths – one is 6 mm long, the other is 4 mm long, both of them being 0.4 mm wide – perpendicular to each other. It is reconstructed using a course grid at the first hierarchical level

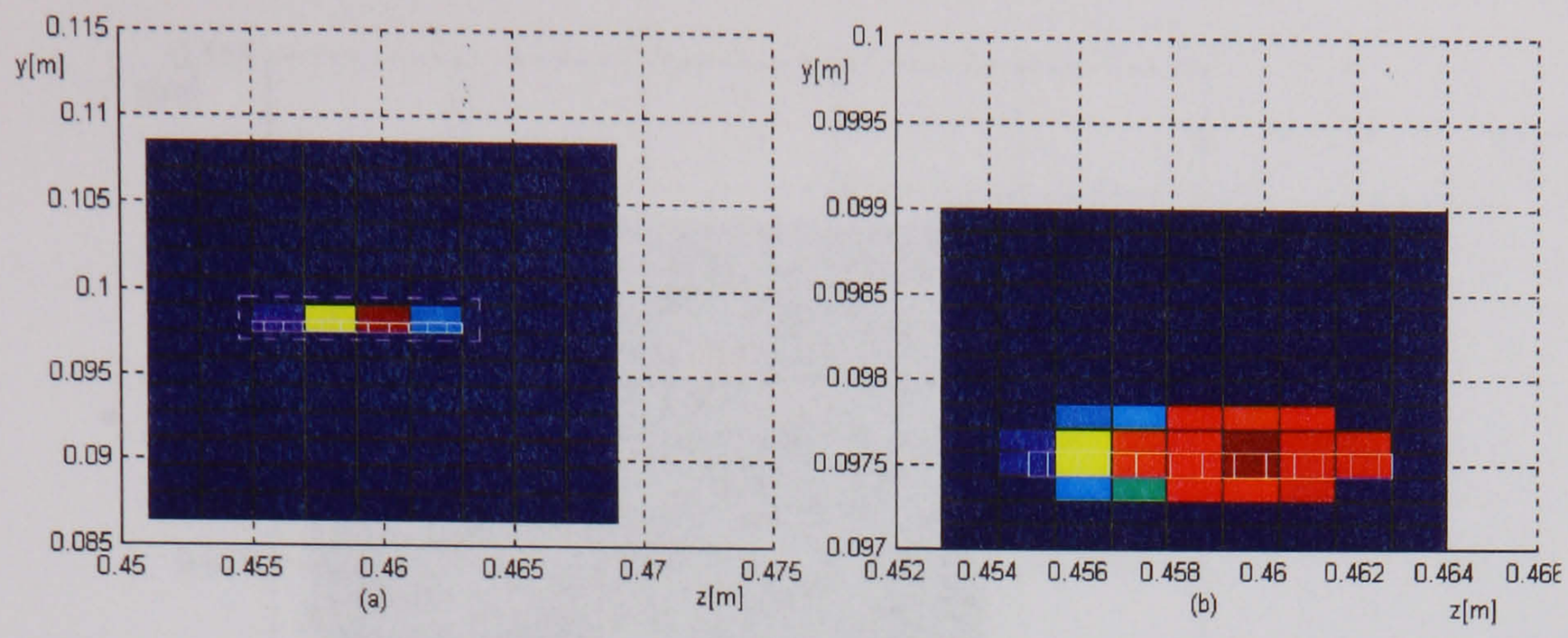
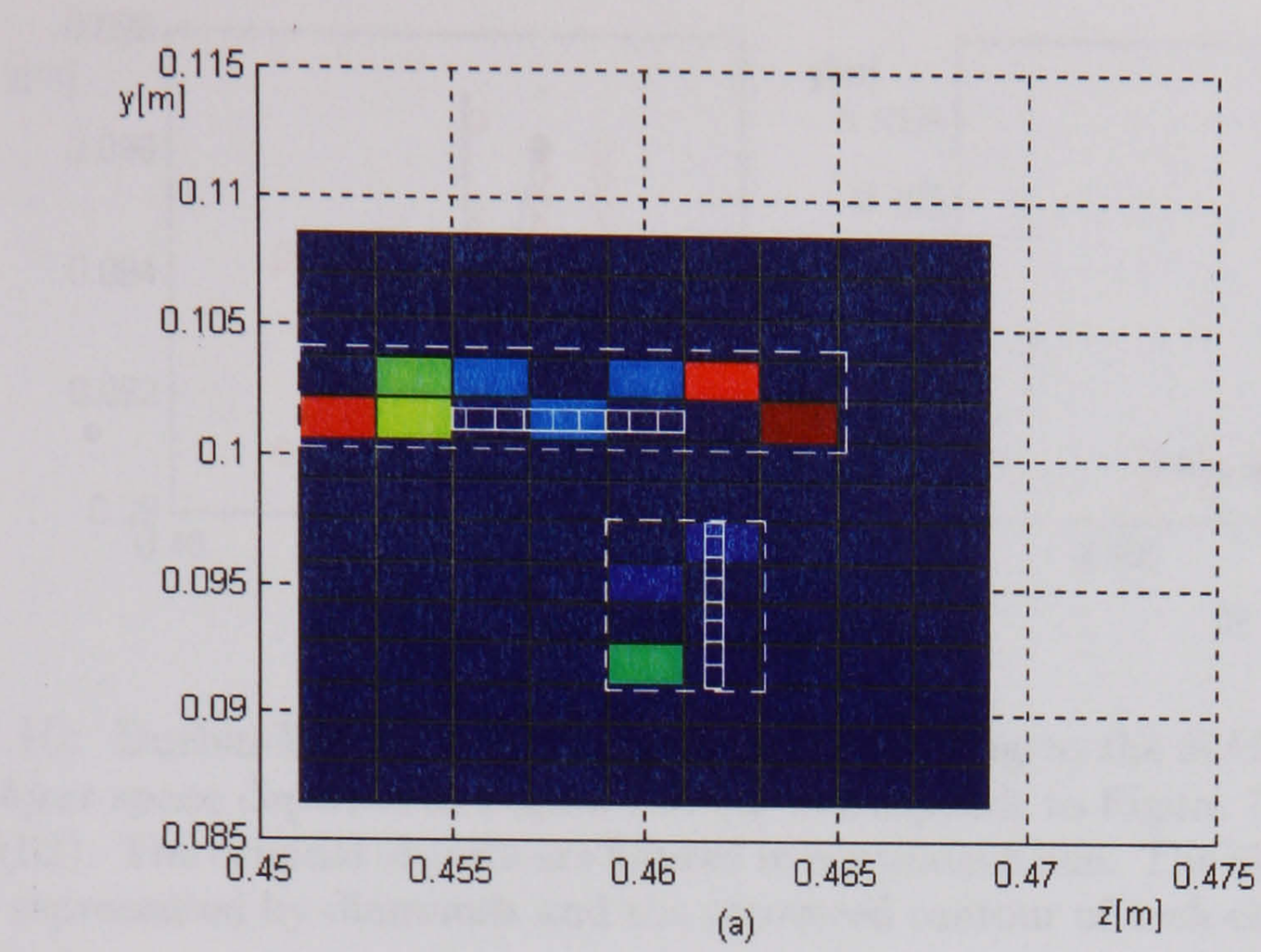


Figure 7.8: Hierarchical AC3D reconstruction of a *face* (plane) orthogonal on the camera aperture from an integral image: (a) initial regularization and high intensity region selection (dashed line); (b) regularization at increased resolution in the region selected at (a). The original object is figured in continuous line.

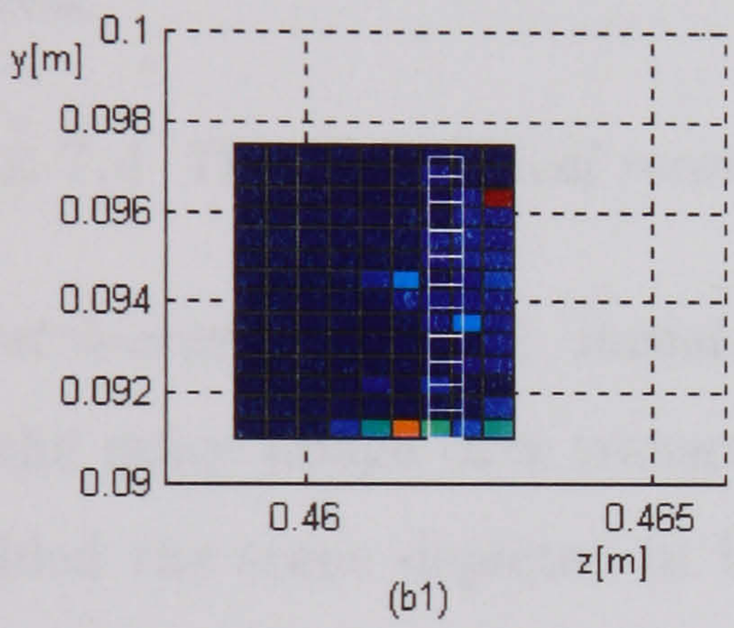
and a finer grid on selected regions at the second one. The integral image is shown in Figure 8.6, Appendix C.

- *First hierarchical level: initial sampling and regularization.* The reconstruction at low resolution is shown in Figure 7.9(a). The objects are correctly positioned and separated in the reconstruction, though their shapes are perturbed. The same as in the single face image case, the face orthogonal to the camera aperture is reproduced with more accuracy than that parallel to the camera aperture. However, it is noticeable that the two objects are clearly separated in the reconstruction. This allows a segmentation of the object space such that the two components can be subsequently searched in parallel, independently, in the high intensity regions marked with dash line.

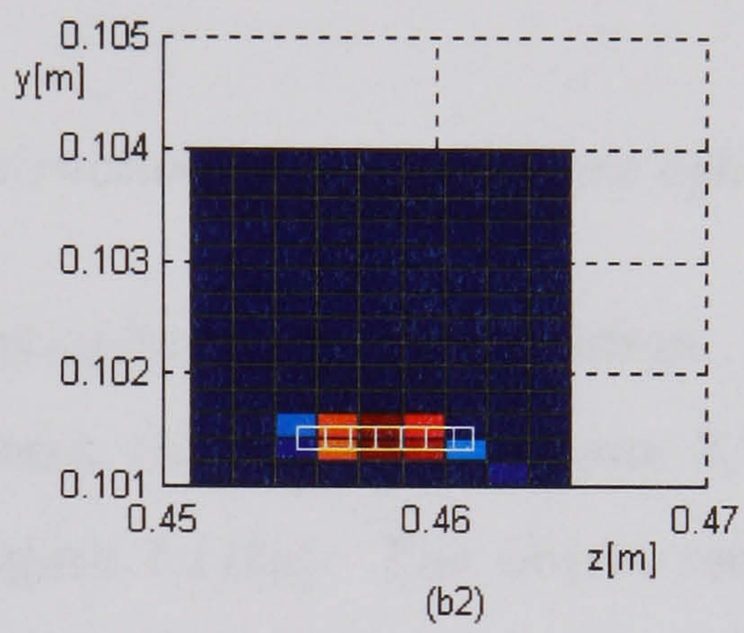
- *Second hierarchical level: localised regularization with increased resolution (algorithm 7.2.2b).* Further processing is performed for the marked regions around the two objects at higher resolution and more accurate reconstructions are obtained (Figures 7.9(b1)-(b2)) using the *multiresolution* approach. The reconstruction obtained for the face orthogonal to the camera aperture is correctly positioned and appears thinned to the width of the object and of correct length. The face parallel to the camera aperture has a relatively more blurred reconstruction, but however thinned and deployed towards the true position, exhibiting better localisation accuracy than the reconstruction obtained at the previous level.



(a)



(b1)



(b2)

Figure 7.9: Hierarchical AC3D reconstruction of an object space composed of *two transparent faces*, one parallel to and another orthogonal on the camera aperture, from an integral image: (a) initial regularization and high intensity region selection (dashed line); (b) regularization at increased resolution in the regions selected at (a). The original objects are figured in continuous line.

- *Depth map drawing.* Finally, the contours of the objects are drawn using two independent Durbin-Willshaw structures. (Figures 7.10(a)-(b)). The contour corresponding to the face parallel to the camera aperture (Figure 7.10(a)) is defined by points – 'stops' – with an average distance to the real object 0.34 mm, whilst the 'stops' that compose the contour corresponding to the orthogonal face (Figure 7.10(b)) are at an average distance of 0.1 mm from the actual object. It is noticed that these points are in the width range of each of the faces, so the final solution is a good approximation of the object scene.

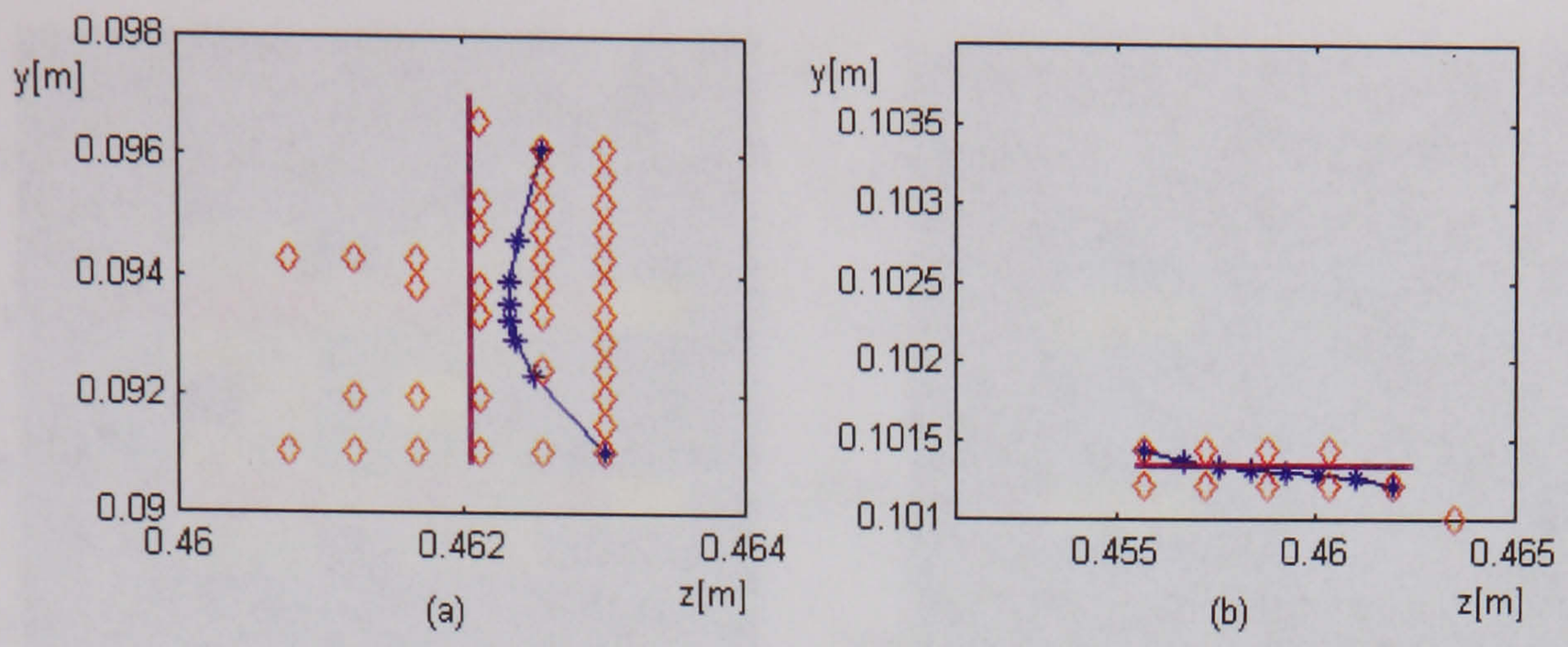


Figure 7.10: Durbin-Willshaw depth maps corresponding to the HAC3D reconstruction of the *two face object* space depicted in Figure 7.9: (a) corresponds to Figure 7.9(b1); (b) corresponds to Figure 7.9(b2). The original objects are figured in continuous line. The high intensity reconstructed points are represented by diamonds and the recovered contour of each object is represented by the starred curves.

EXAMPLE 7.4 *The hierarchical reconstruction of a transparent cylinder*

- *First hierarchical level: initial sampling and regularization.* The processing of the noisy image of a transparent thin cylinder (Figure 8.7, Appendix C) yielded the scene depicted in Figure 7.11(a). The object recovered presents lateral (y) symmetry and is located around the true position of the object, but the shape of the object is not clear. This is not unexpected, as the sampling grid is very sparse in the object space and cannot retain the curvature of the object's contour at this resolution. However, the high intensity region around the object is selected for further regularization.

- *Second hierarchical level – local regularization at the same resolution (algorithm 7.2.2a).* For better accuracy, *local regularization* is performed in the vicinity (dashed line in Figure 7.11(a)) of the high intensity points. The result is shown in Figure 7.11(b). The new reconstruction contains less high intensity points than the first one. It is noticed that it suggests a hollowed contour, the 'front' of the object (left hand side in the picture) being more prominent.

- *Depth map drawing.* In order to link the resulted 'edges' and produce the contour, the weighted Durbin–Willshaw scheme described in section 7.3 has been used. The curve fitting the intensity map from Figure 7.11(b) is depicted in Figure 7.12(a). It matches very well the actual contour of the object, shown in continuous line on the same figure, the average distance of a 'stop' on

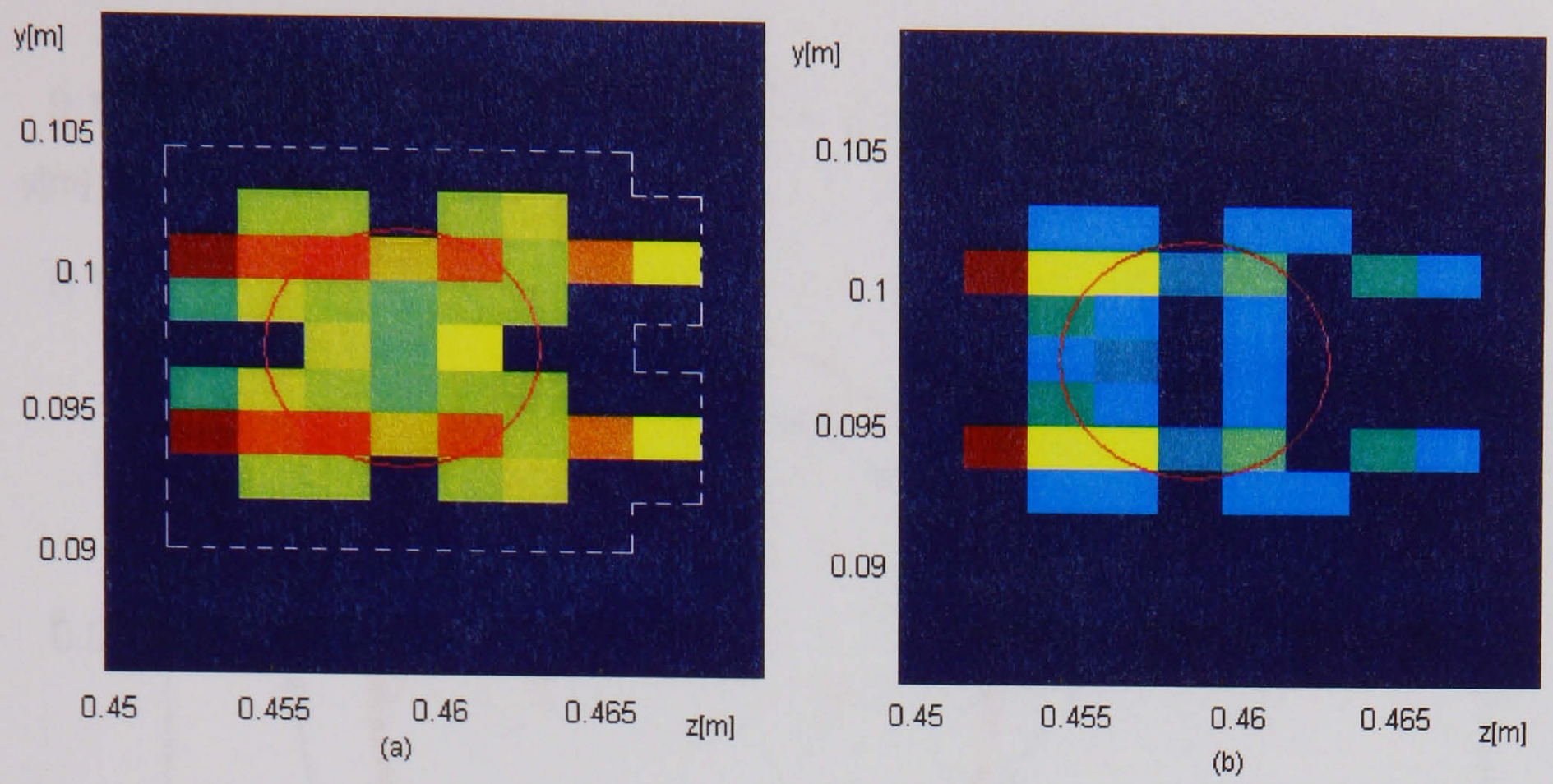


Figure 7.11: Hierarchical AC3D reconstruction of a *transparent cylinder* from a noisy integral image: (a) initial regularization and high intensity region selection (dashed line); (b) localised regularization at the same resolution in the region selected at (a). The original object is figured in continuous line.

the contour to the real object being 0.55 mm. Figure 7.11(b) depicts three dimensionally the reconstruction of the cylinder.

EXAMPLE 7.5 *The hierarchical reconstruction of a an opaque cube*

A few variants for the reconstruction of the object space from a noisy integral image of an opaque cube (Figure 8.8, Appendix C) have been considered.

Variant 1.

- *First hierarchical level: initial sampling and regularization.* The processing of the image of the visible sides of an opaque cube yielded the scene depicted in Figure 7.13. The object recovered at this stage is symmetric and the two sides are clearly separated. The depth and lateral position are correct. The high intensity area is selected and used at the following step.

- *Second hierarchical level – local regularization at increased resolution (algorithm 7.2.2b).* For better accuracy, *local regularization* is performed in the previously selected vicinity of the high intensity points at increased resolution. The result is shown in Figure 7.14. It can be noticed that the final reconstruction follows the actual shape of the object, but is blurred and slightly larger than the object. For a more computationally economic

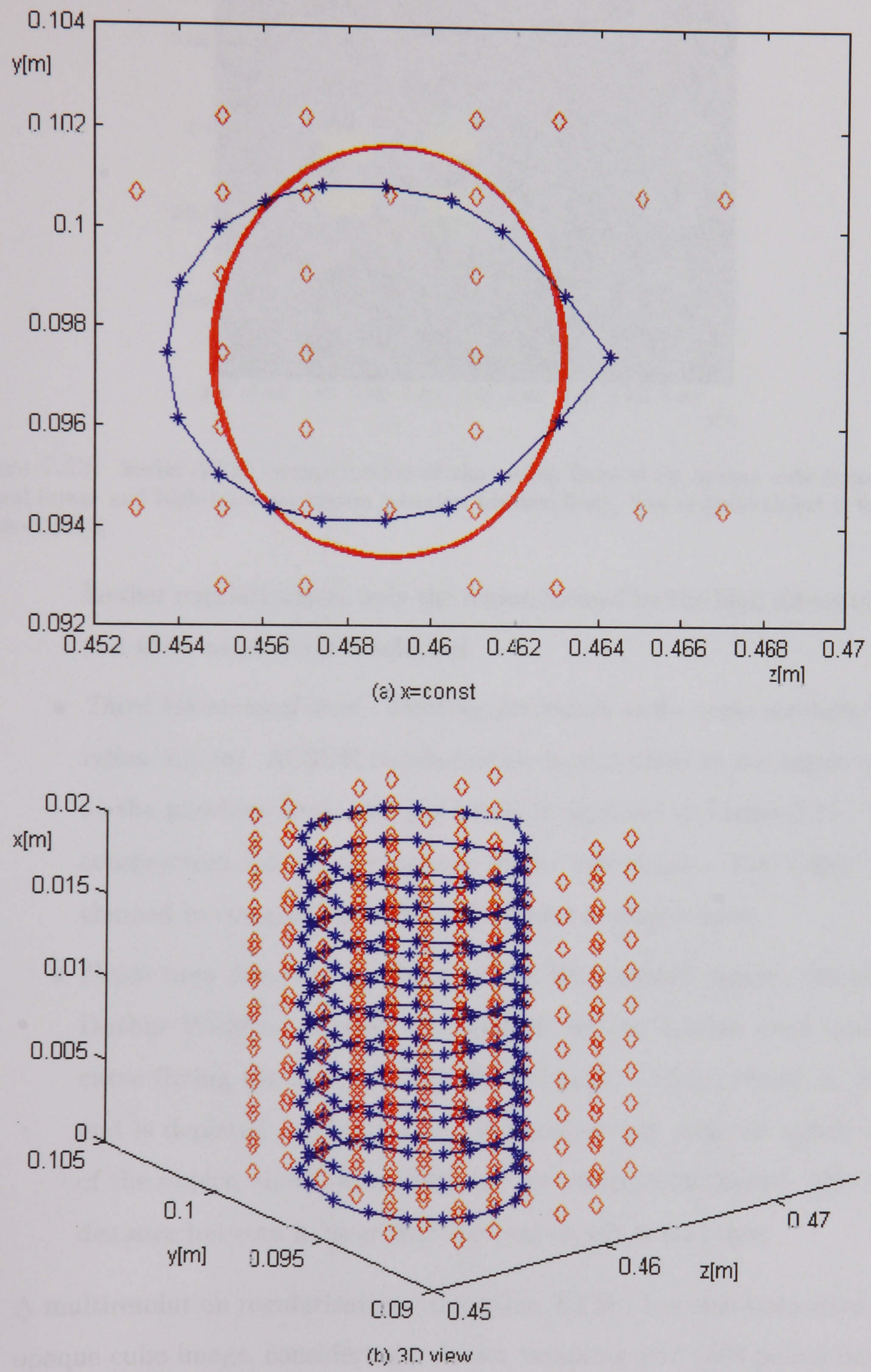


Figure 7.12: Durbin-Willshaw depth map corresponding to the HAC3D reconstruction of a *transparent cylinder*, depicted in Figure 7.11: (a) Transversal section with a plane $x = const$. The original object is figured in continuous line. (b) 3D view of the reconstruction. The high intensity reconstructed points are represented by diamonds and the recovered contour of the object is represented by the starred curve(s).

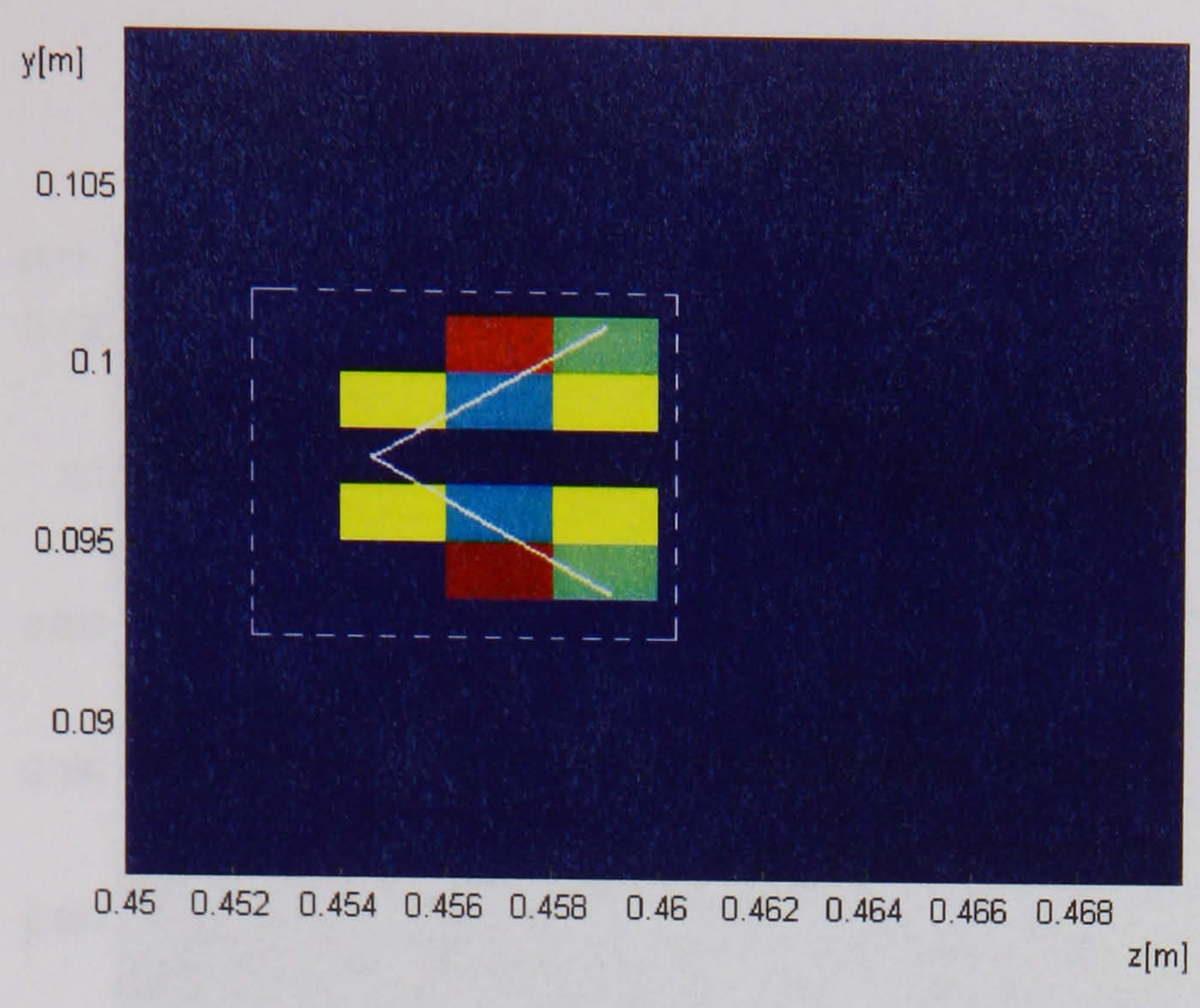


Figure 7.13: Initial AC3D reconstruction of the visible faces of an *opaque cube* from a *noisy* integral image and high intensity region selection (dashed line). The original object is figured in continuous line.

further regularization, only the region formed by the high intensity points and their neighbours is selected.

- *Third hierarchical level – local regularization at the same resolution (algorithm 7.2.2a).* AC3DR regularization is performed in the region selected at the previous level, and the result is depicted in Figure 7.15. The reconstruction follows the contour of the real object. The 'edges' appear thinned in comparison with the previous reconstruction.

- *Depth map drawing.* In order to link the resulted 'edges', the weighted Durbin–Willshaw scheme described in section 7.3 has been used. The curve fitting the intensity map from Figure 7.15 is defined by 18 stops and is depicted in Figure 7.16. It matches very well the actual contour of the object, shown in continuous line on the same figure. The average distance between a 'stop' and the real object is 0.24 mm.

Variant 2. A multiresolution regularization (algorithm 7.2.2b) has also been tried for the opaque cube image, considering a denser sampling grid (600 points instead of 136 points) in the area surrounding the high intensity region obtained at the first level reconstruction presented in Figure 7.13. The whole processing is summarised in the following:

- *First hierarchical level: initial sampling and regularization.* This step

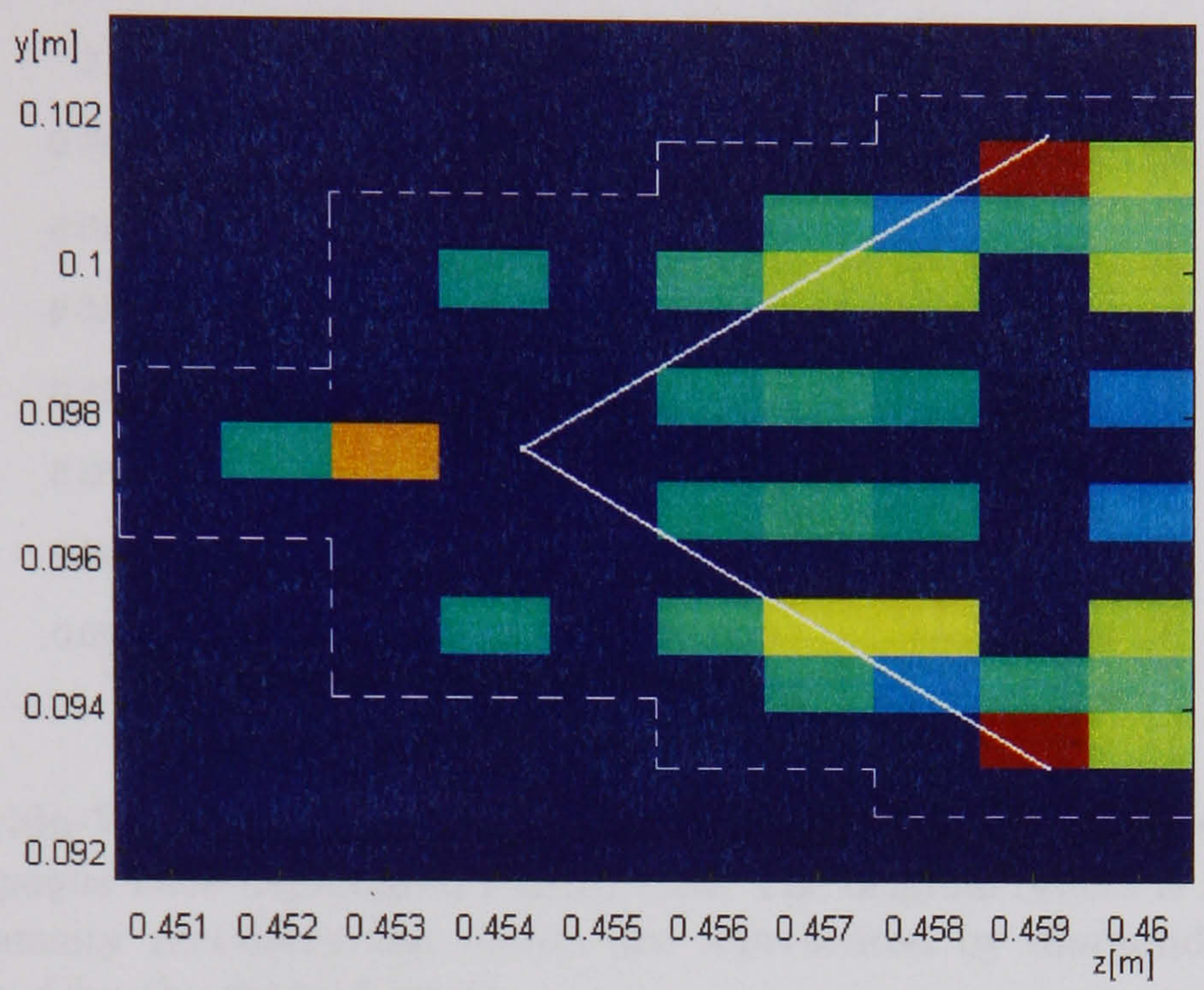


Figure 7.14: Second hierarchical level AC3D reconstruction of the visible faces of an *opaque cube* from a *noisy* integral image in the region selected at the first phase (Figure 7.13) using higher resolution. High intensity region selection (dashed line). The original object is figured in continuous line.

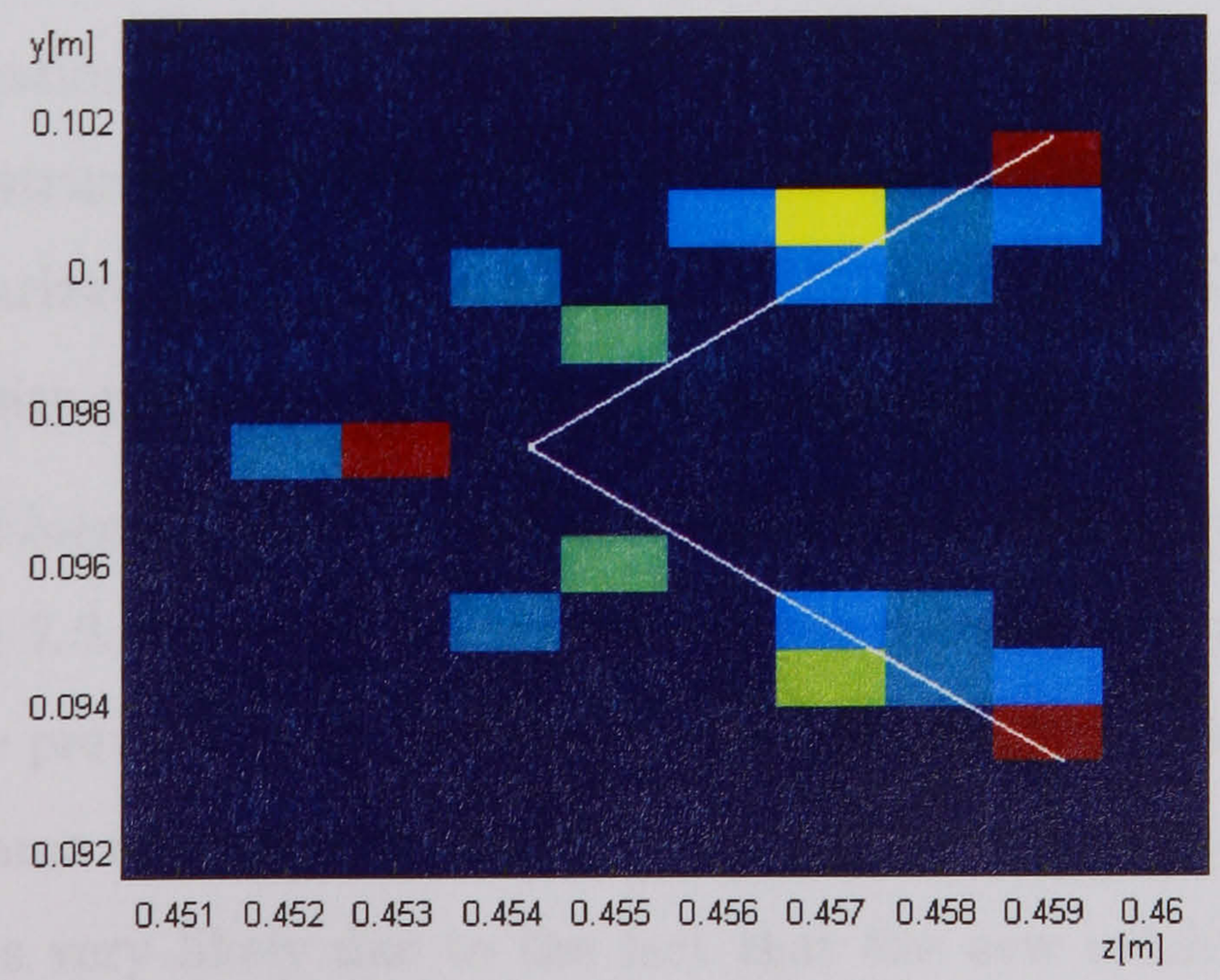


Figure 7.15: Third hierarchical level AC3D reconstruction of the visible faces of an *opaque cube* from a *noisy* integral image in the region selected at the second phase (Figure 7.14) using the same resolution. The original object is figured in continuous line.

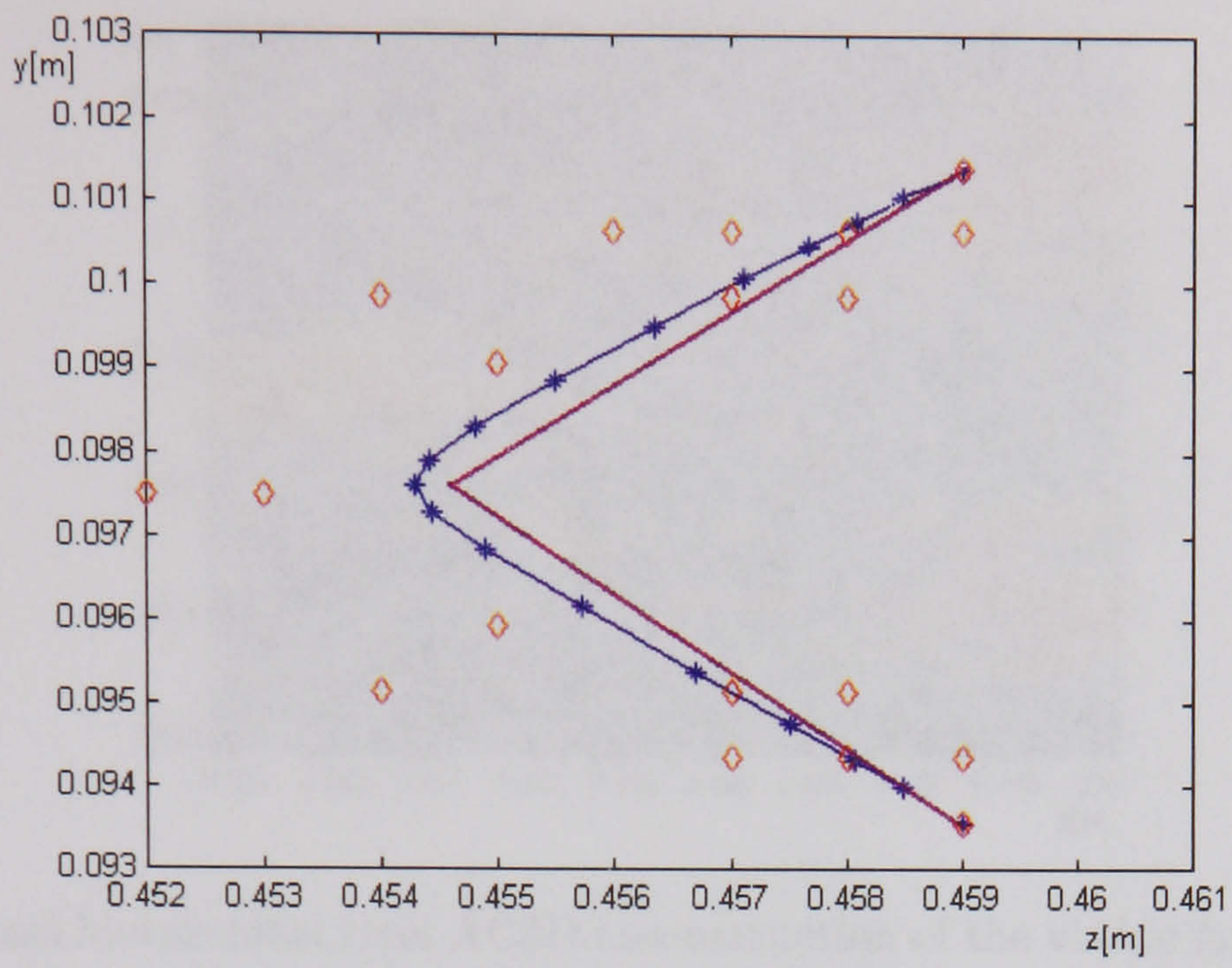


Figure 7.16: Durbin-Willshaw depth map corresponding to the HAC3D reconstruction of the visible faces of an *opaque cube* depicted in Figure 7.15. The original object is figured in continuous line. The high intensity reconstructed points are represented by diamonds and the recovered contour is represented by the starred curve.

corresponds to the result shown in Figure 7.13 and already discussed at Variant 1.

- *Second hierarchical level – local regularization at increased resolution (algorithm 7.2.2b).* The region selected at the previous step has been resampled at a resolution more than 4 times higher than the one used at the second level regularization in Variant 1. The multiresolution regularization result is shown in Figure 7.17. It can be noticed that the reconstruction is more blurred than the one obtained previously by local regularization in Variant 1 (Figure 7.14). The region formed by the high intensity points and their neighbours is selected.

- *Third hierarchical level – local regularization at the same resolution (algorithm 7.2.2a).* AC3DR regularization is performed in the region selected at the previous level, and the result is depicted in Figure 7.18. It exhibits the same accentuated blur, in spite of the repeated regularization. This blur is very likely due to the fact that the new resolution of the object space exceeds the capability of the optical system.

- *Depth map drawing.* To link the resulted 'edges', a weighted Durbin–Willshaw scheme with 18 'stops' has been used. The curve fitting the intensity map from Figure 7.18 is depicted in Figure 7.19. Clearly, the

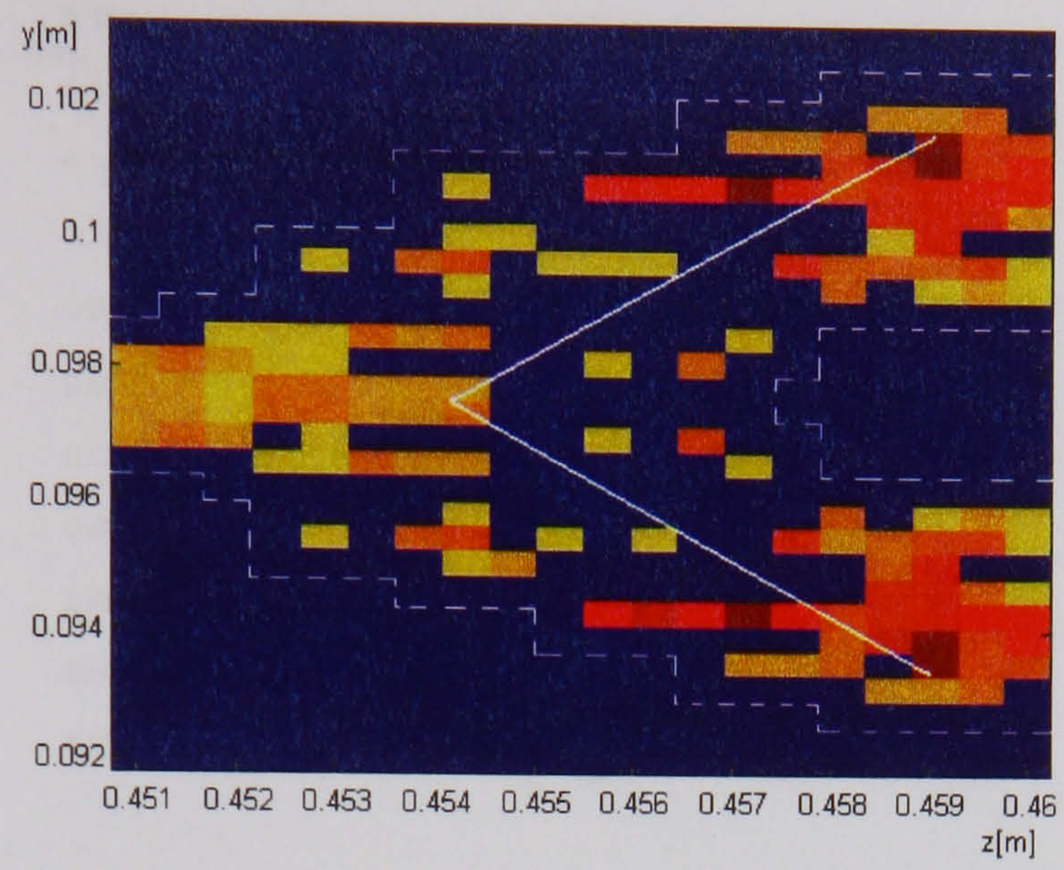


Figure 7.17: Second hierarchical level AC3D reconstruction of the visible faces of an *opaque cube* from a *noisy* integral image in the region selected at the first phase (Figure 7.13) using a resolution 4 times higher than that used in Figure 7.14. High intensity region selection (dashed line). The original object is figured in continuous line.

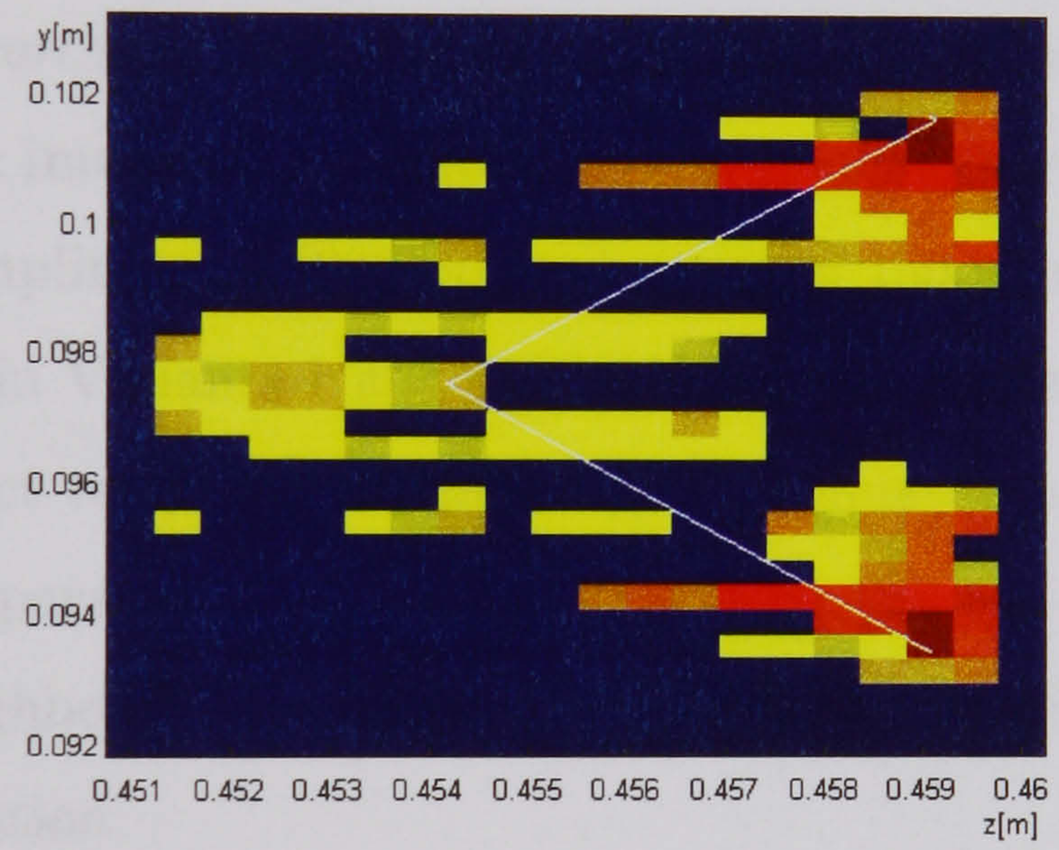


Figure 7.18: Third hierarchical level AC3D reconstruction of the visible faces of *an opaque cube* from a *noisy* integral image in the region selected at the second phase (Figure 7.17) using the same resolution. The original object is figured in continuous line.

match is not as good as the one in Figure 7.16, but it follows the actual contour of the object, shown in continuous line on the same figure. The average distance between a 'stop' and the real object is 0.6 mm.

In conclusion, when the resolution of the reconstructed object space is increased to a level where the resolution capabilities of the optical system are exceeded, the reconstruction thus obtained is very blurred. However, this lack of definition in the reconstruction is overcome by the robustness of the weighted Durbin-Willshaw scheme, which is able to recover to a good extent the actual contour of the object even from a very blurred intensity map.

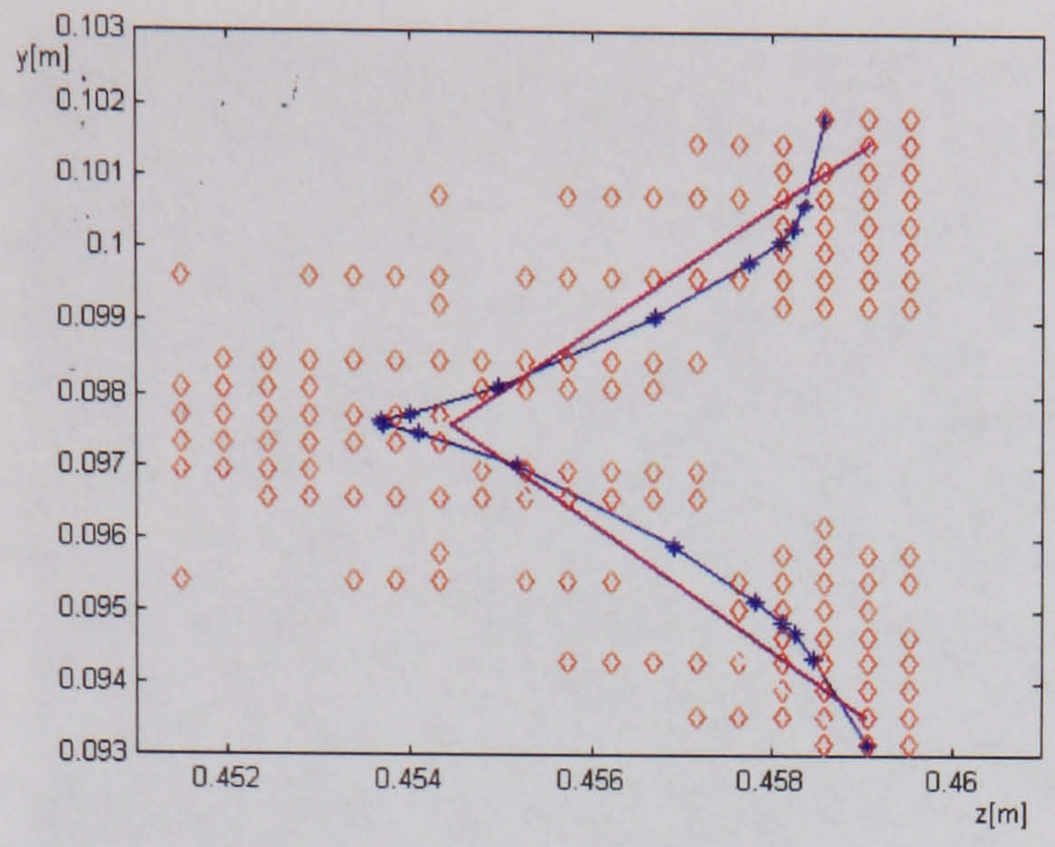


Figure 7.19: Durbin-Willshaw depth map corresponding to the hierarchical AC3D reconstruction of the visible faces of an *opaque cube* depicted in Figure 7.18. The original object is figured in continuous line. The high intensity reconstructed points are represented by diamonds and the recovered contour is represented by the starred curve.

Variant 3.

- *First hierarchical level: initial sampling and regularization.* The processing of the image of the visible sides of an opaque cube using a relatively dense sampling grid (600 points instead of 150 points at the corresponding level in Variants 1 and 2) yielded the scene depicted in Figure 7.20. The object recovered at this stage is symmetric and the two sides are clearly separated. The region formed by the high intensity points and their neighbours is selected for a computationally inexpensive further regularization.

- *Second hierarchical level – local regularization at the same resolution (algorithm 7.2.2a).* Local regularization is performed in the vicinity of the high intensity points selected at the previous step. The result is shown in Figure 7.21. It can be noticed that this reconstruction follows the actual contour of the object.

- *Depth map drawing.* In order to link the resulted 'edges', the weighted Durbin–Willshaw scheme described in section 7.3 has been used. The curve fitting the intensity map from Figure 7.20 is depicted in Figure 7.22. It matches very well the actual contour of the object, shown in continuous line on the same figure, the average distance between a 'stop' and the real object being 0.09 mm. It is the most accurate reconstruction obtained for the visible faces of the opaque cube. A three dimensional

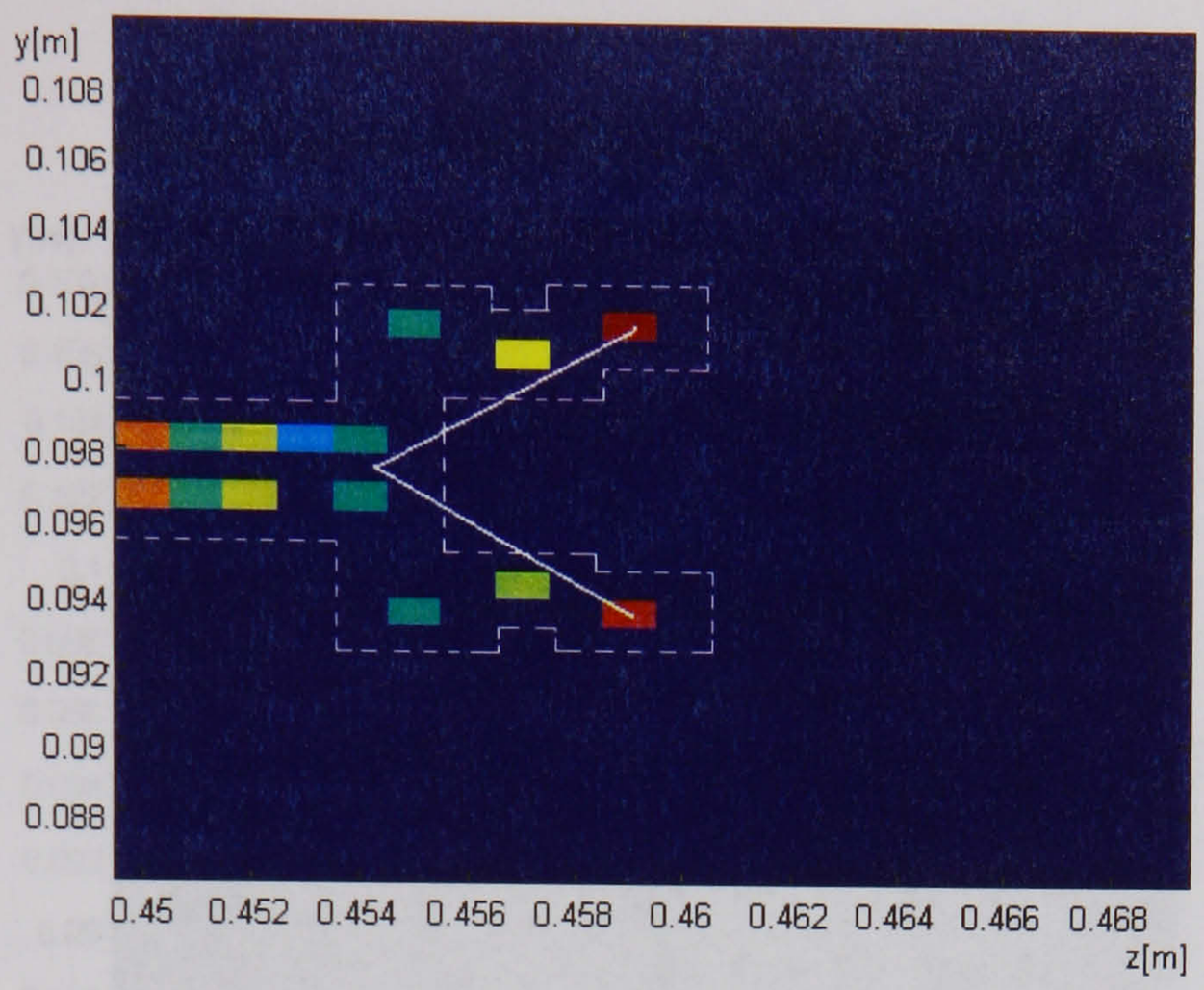


Figure 7.20: Initial AC3D reconstruction of the visible faces of an *opaque cube* from a *noisy* integral image at high resolution. The high intensity region is selected (dashed line). The original object is figured in continuous line.

view of the reconstructed object is presented in Figure 7.23.

7.5 Conclusions

The chapter has presented a novel solution to the object reconstruction and depth extraction from 3D-integral images as an inverse problem. This is a hierarchical adaptive constrained 3D-reconstruction algorithm combined with a depth mapping technique. The regularization involved by the algorithm is based on the use of sigmoid functions to determine a positive bounded constrained least squares solution which provides information about the number of objects in the scene, their shape and absolute and relative position. This scheme can be hierarchically reapplied to the relevant regions of the object space in order to produce higher resolution and higher accuracy intensity maps of the scene. Finally, a weighted Durbin-Willshaw elastic net is employed to produce a depth map of the object space from the intensity map. The simulation results have shown that the number of objects, absolute and relative positions and shape are recovered accurately. The hierarchical scheme is very robust to noise, produces high resolution reconstructions and is computationally efficient, as only the relevant regions of the scene are oversampled. The depth mapping procedure is very robust to noise and interpolation perturbations, overcoming the inherent limitations of the discrete data processing.

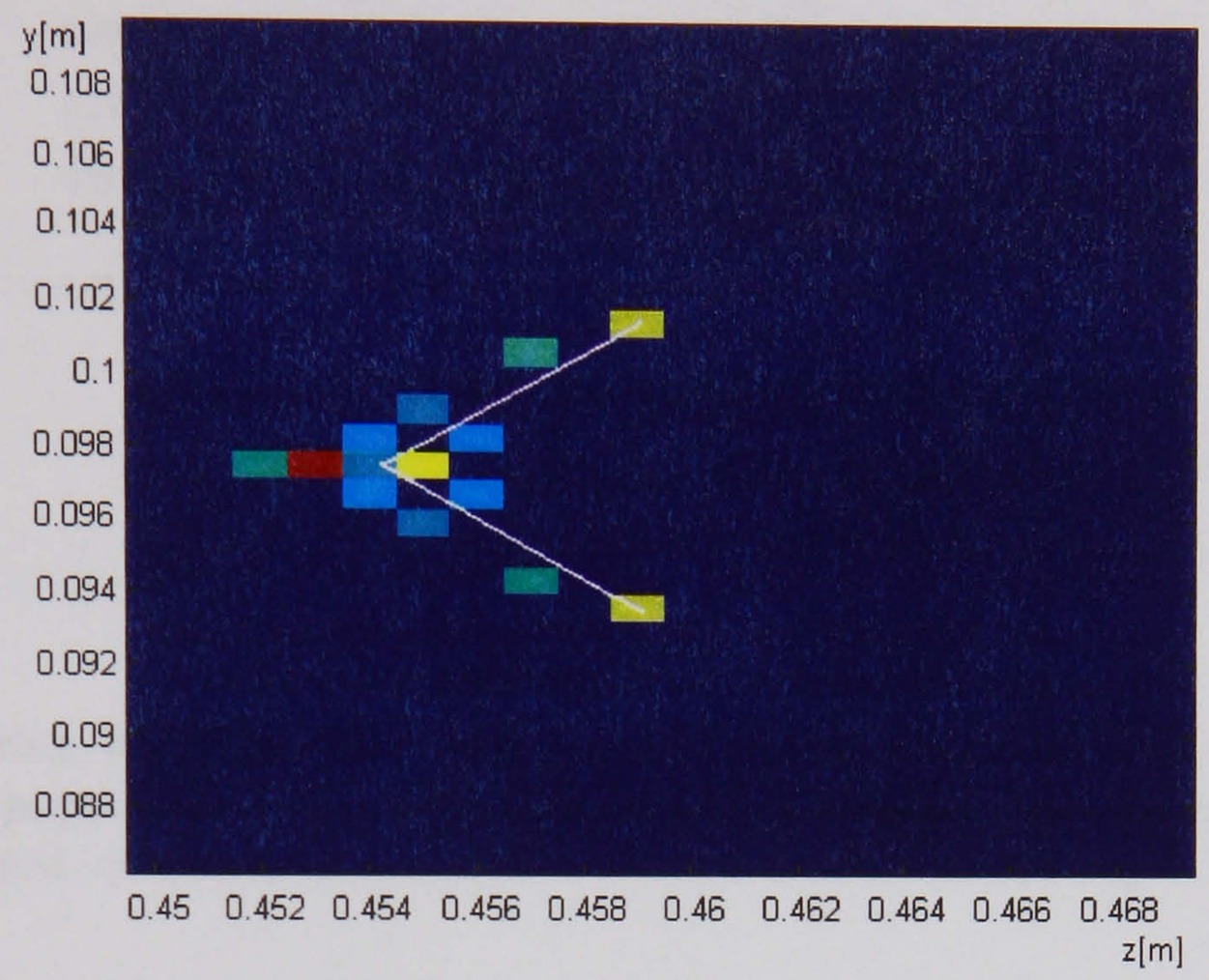


Figure 7.21: Second hierarchical level AC3D reconstruction of the visible faces of an *opaque cube* from a *noisy* integral image in the region selected at the first phase (Figure 7.20) using the same resolution. The original object is figured in continuous line.

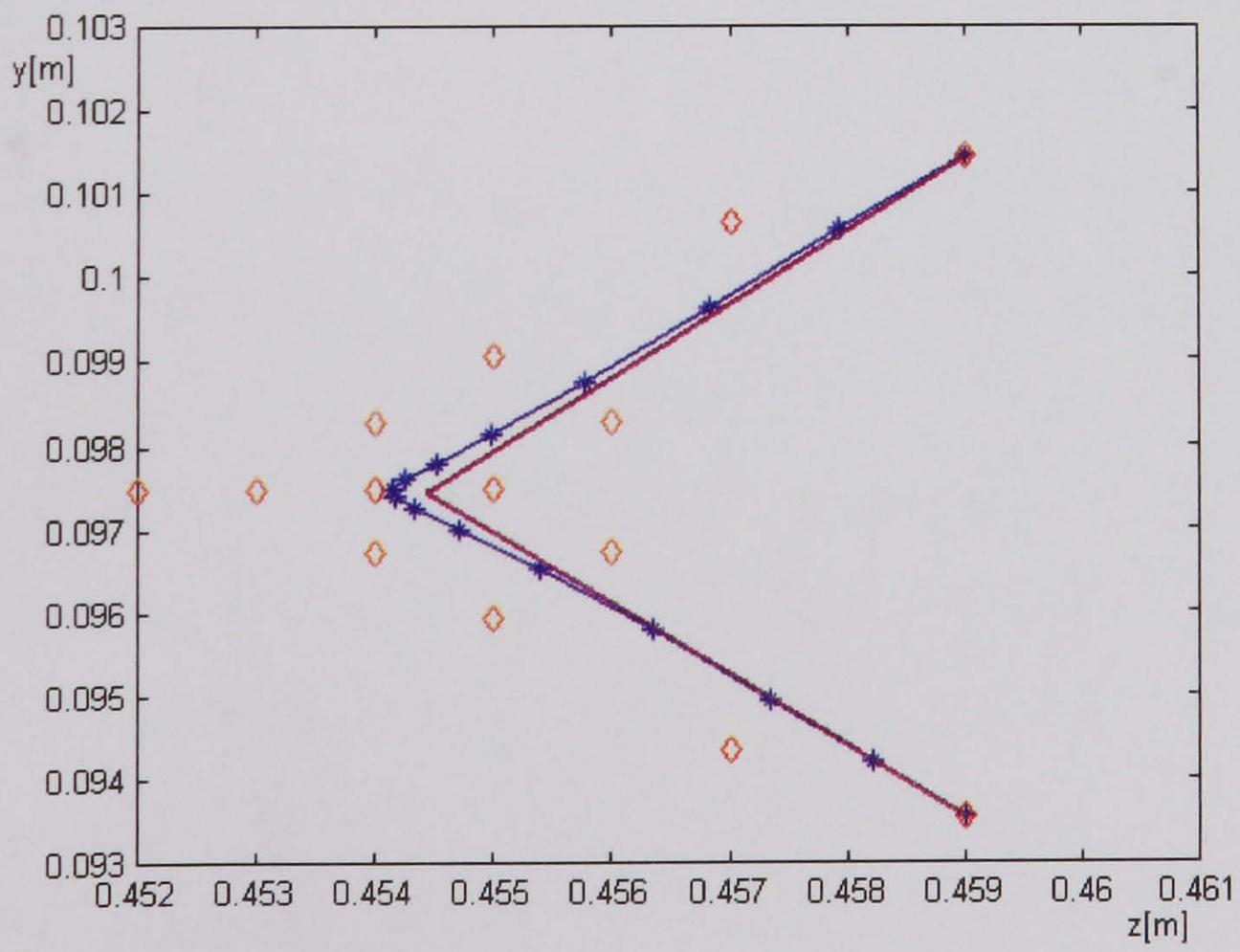


Figure 7.22: Durbin-Willshaw depth map corresponding to the hierarchical AC3D reconstruction of the visible faces of an *opaque cube* depicted in Figure 7.21 (transversal section with a plane $x = const$). The high intensity reconstructed points are represented by diamonds and the recovered contour of the object is represented by the starred curve. The original object is figured in continuous line.

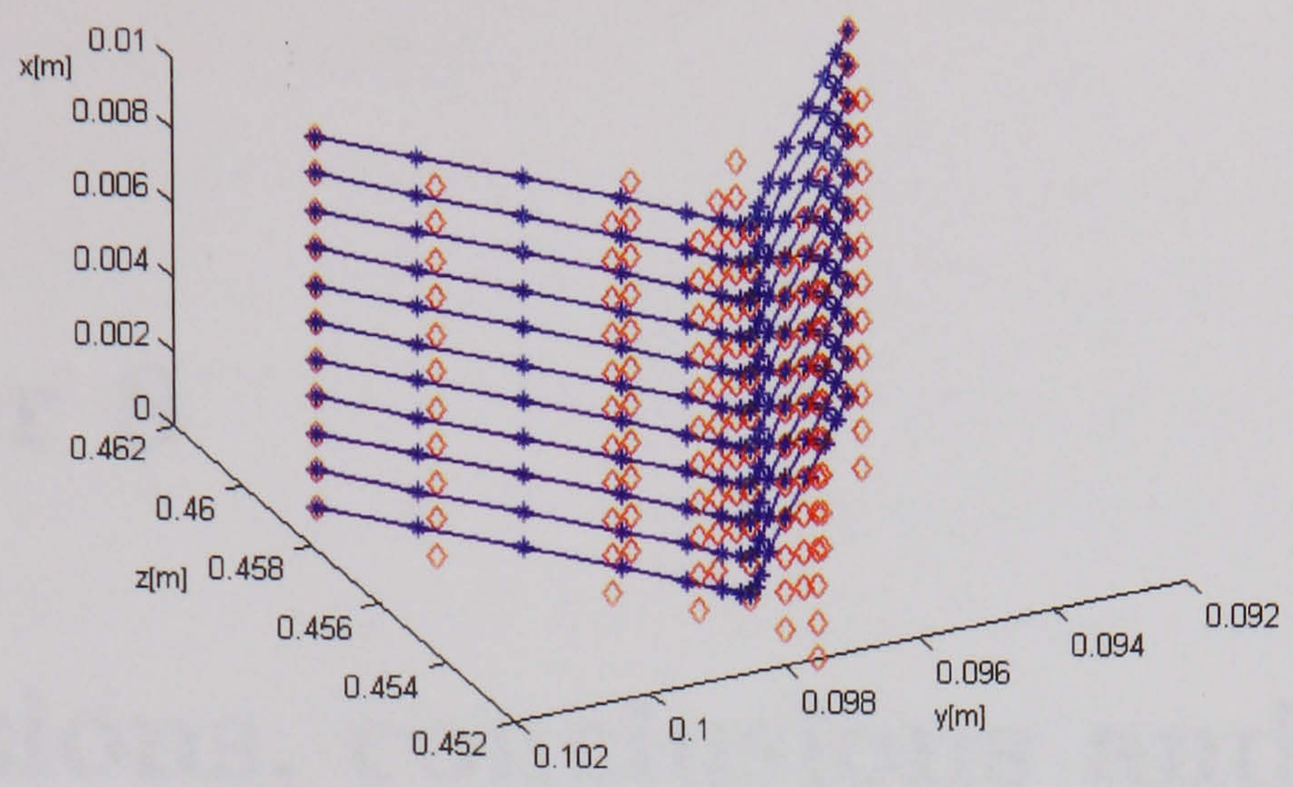


Figure 7.23: Durbin-Willshaw depth map corresponding to the HAC3D reconstruction of the visible faces of an *opaque cube* depicted in Figure 7.21 (3D view). The high intensity reconstructed points are represented by diamonds and the recovered contour of the object is represented by the starred curves.

The applicability of the algorithm is not confined to integral imaging. The overall idea can be also used in holography and multi-view systems, with appropriate imaging operators. AC3DR provides solutions for any inverse problems whose sought solutions are positive and bounded. Also, the hierarchical multiresolution strategy and the modified Durbin-Willshaw scheme can be used independently in conjunction with any of the existing regularization methods ([60]).

Chapter 8

Discussions, conclusions and further work

8.1 Discussion and conclusions

The present research has approached integral imaging from the perspective of decoding the depth information stored in a planar recording of a 3D-integral image.

The operation of the 3D integral camera system and the associated image formation and recording processes have been presented. A mathematical model of the geometry and physics of the optical process involved by the 3D unidirectional integral recording camera system, which uses cylindrical surface microlenses to produce images with continuous horizontal parallax, has been described. The coordinates of the intermediary, integrated, and recorded data have been computed, and an analysis of the imaging field corresponding to a microlens in the recording array has been proposed, using a pinhole approximation hypothesis and a 2D-refraction model. The point spread function of the entire process and its dependence on point depth has been analysed. The problem of computing the number of recorded micro-intensity distributions and the dependence on point position has been considered. The results illustrate the dependence of recorded intensity distributions and object footprint on spatial position. The model has been generalised to the omnidirectional integral camera system, which produces integral images with continuous parallax in all directions. Three variants of the camera, using square, circular, or hexagonal based microlenses, have been considered and compared in terms of retained angu-

lar information and point spread. Two types of microlens packaging in the array (rectangular net and honeycomb) have been analysed, and the number of recorded micro-intensity distributions for a single physical point has been evaluated for each of the packaging types, in relation to the lateral position and depth of the point. The analysis has taken into account the combined effect of all the microlenses of the arrays in the image formation, and has shown that the recording process using honeycomb packed circular or hexagonal microlenses retains more angular details than that corresponding to a rectangularly packed square or circular based microlens array. Moreover, the hexagonal microlens arrays do not have inactive areas and ensure an efficient use of the recording surface.

The point spread function of the entire process and its dependence on point depth has been presented for square, circular, and hexagonal based microlens configurations in both close and remote imaging conditions. The use of a camera based on square microlenses is advisable when recording scenes which are reconstructed at a distance away from the recording screen, while circular based microlenses are recommended for recording scenes which straddle the recording screen. The hexagonal based microlens system presents a good trade-off between spread performance at both remote and close imaging and the retained angular information. It is therefore recommended when a wide depth range is recorded.

The mathematical models have provided an understanding of the way depth information is encoded in the planar recordings of 3D-integral images and have opened the way for a new formulation of the depth estimation problem. The inverse process of image formation allows the recovery of an intensity map of the object space and a subsequent estimation of the object depth. Thus, depth extraction has been associated with object space reconstruction and formulated as an inverse problem of the imaging process. This has proven very ill-conditioned. To overcome the ill-conditioning, regularization methods have been used to obtain physically meaningful solutions of the problem. An extensive overview of existing regularization methods and of their convergence properties has been presented.

Simulation results have proven that the methods provide very accurate intensity maps of the object space when some of the sampling points are at or very near the actual object points, but much poorer results when the sampling grid is sparse

around the object. The regularized solutions obtained are generally realistic in terms of their depth extraction capability, but the final solutions generally contain negative values, and the computational cost required by an accurate reconstruction is high. Interpolation of the real intensity values at the sampling points produces high perturbation of the results and makes the assessment of the solutions difficult and inherently ambiguous. Moreover, the resolution of the reconstructions is drastically confined by the number of pixels in the image.

To overcome the difficulties posed by the existence of negative values in the solutions, interpolation and low resolution problems, a new hierarchical depth extraction algorithm has been proposed. The regularization involved by the algorithm (AC3DR) is based on the use of sigmoid functions to determine a bounded constrained least squares solution which provides information about the number of objects in the scene, their shape and absolute and relative position. This scheme is hierarchically reapplied to the relevant regions of the object space in order to produce higher resolution and higher accuracy intensity maps of the scene. Finally, a weighted Durbin-Willshaw elastic net is employed to produce a depth map of the object space from the intensity map. Simulation results have shown that the number of objects, absolute and relative positions and shape are accurately recovered. The hierarchical multiresolution strategy is very robust to noise, produces high resolution reconstructions and is computationally efficient, as only the relevant regions of the scene are oversampled. The depth mapping procedure is very robust to noise and interpolation perturbations, overcoming the inherent limitations of the discrete data processing.

The applicability of the algorithm is not confined to integral imaging. The overall idea can be used in holography and multi-view system, with appropriate imaging operators. AC3DR provides solutions for any inverse problem whose sought solutions are positive and bounded. Also, the hierarchical multiresolution strategy and the modified Durbin-Willshaw scheme can be used independently in conjunction with any of the existing regularization methods.

8.2 Further work

The research work presented in this thesis can be enhanced through further theoretical developments and practical implementations.

Possible theoretical refinements of the Hierarchical Depth Extraction Algorithm may be achieved through:

- A parametric study for the AC3DR algorithm aimed at deriving the optimal relationship between the sigmoid parameter σ and the threshold T. The practical outcome should be an automatic choice of the optimal parameters for a given picture and sampling grid.

- A study of the global convergence of the method.

- A study of the filtering properties of the sigmoid. The sigmoid can be used as a thresholding function instead of the standard convex projection operator in projected Landweber type methods. Thus, new regularization methods can be derived.

- A study of the relationship between the resolution capabilities of the camera system and the optimal resolution of the reconstruction. Possible sampling optimisation as outcome.

- A parametric study of the Durbin-Willshaw depth mapping scheme. The aim should be to derive the optimal number of 'stops' that define the final contour from a given set of reconstructed high intensity points. An optimal parameter choice for the radius of attraction of a 'city' and for the parameter α that governs the mutual attraction between two 'stops' is also useful.

- Depth mapping using other self organising feature maps known from unsupervised learning neural network techniques, e.g. Kohonen algorithms, 'snake'.

Further implementations and testing of the existing algorithm may concentrate on:

- Tests performed on real images.

- Implementation of the method for omnidirectional integral images.

- Development of a software tool able to automatically select regularization parameters, execute hierarchical and localized regularization, and derive the depth map. Computational difficulties may be reduced through a parallel processing approach.

A long term continuation of the project may concentrate on merging real and/or computer generated integral 3D objects into an integral image and a possible development of an 'integral virtual studio'.

Appendix A: Publications

Journals:

Silvia Manolache, A. Aggoun, M. McCormick, N. Davies, S.Y. Kung, **Analytical model of a three-dimensional integral image recording system that uses circular and hexagonal based spherical surface microlenses**, Journal of the Optical Society of America A, vol. 18, no. 8, pp. 1814-1821, August 2001.

Silvia Manolache, S.Y. Kung, M. McCormick, A. Aggoun, **3D-object space reconstruction from planar recorded data of 3D-integral images**, to appear in Journal of VLSI Signal Processing Systems, Kluwer Acad. Publ., 2001.

Conferences:

Silvia Manolache, M. McCormick, S.Y. Kung, **Hierarchical adaptive regularization method for depth extraction from planar recdording of 3D-integral images**, Proceedings of IEEE International Conference for Acoustics, Speech and Signal Processing (ICASSP), vol. 3, pp. III-1433-III-1436, Salt Lake City, May 2001.

Silvia Manolache, S.Y. Kung, M. McCormick, A. Aggoun, **Regularization methods for object space reconstruction from 3D-integral images**. Actes des Conferences Numerisation 3D, paper 12, Paris, April 2001.

Silvia Manolache, A. Aggoun, M. McCormick, N. Davies, **A mathematical model of a 3D-lenticular integral recording system**, Proceedings of IEEE Vision, Modeling and Visualization Conference, Erlangen, pp. 51-58, 1999.

Appendix B: Computations for the PSF of a hexagonal microlens

The amplitude of the light wave in the focal plane is given by:

$$\mathcal{A}_{hex}(X,Y) = C \exp\left(-\frac{ik}{2f}(X^2+Y^2)\right) \mathcal{F}(\psi_1) \star F(\psi_2)\,, \tag{8.1}$$

where $\psi_1 = \exp\left(-(ik\lambda^2 f^2/(2a))(\hat{x}^2+\hat{y}^2)\right)$, therefore, from the definition of the Fourier transform:

$$\mathcal{F}(\psi_1) = \frac{a}{\lambda f^2} \exp\left(-\frac{i\pi a}{\lambda f^2}(X^2+Y^2)\right)\,, \tag{8.2}$$

and ψ_2 represents the Fourier transform of the characteristic function of a hexagonal domain of side $\hat{l} = l/\lambda f$, f being the focal distance of a microlens ([57]):

$$\psi_2(\hat{x},\hat{y}) = \begin{cases} 1\,, & \text{if } (\hat{x},\hat{y}) \in Hex(\hat{l}) \\ 0\,, & \text{otherwise}\,. \end{cases} \tag{8.3}$$

$\mathcal{F}(\psi_2)$ can be decomposed into two integrals by dividing the hexagonal domain (see Figure 8.1):

$$\begin{aligned} \mathcal{F}(\psi_2)(X,Y) &= \int_{-\frac{\hat{l}\sqrt{3}}{2}}^{0} \left(\int_{-\hat{l}-\frac{\sqrt{3}}{3}X}^{\hat{l}+\frac{\sqrt{3}}{3}X} \exp(-i2\pi(X\hat{x}+Y\hat{y})\,dy \right) dx \\ &\quad + \int_{0}^{\frac{\hat{l}\sqrt{3}}{2}} \left(\int_{-\hat{l}+\frac{\sqrt{3}}{3}X}^{\hat{l}-\frac{\sqrt{3}}{3}X} \exp(-i2\pi(X\hat{x}+Y\hat{y})\,dy \right) dx \\ &= I_1 + I_2\,. \end{aligned} \tag{8.4}$$

The first integral yields:

$$\begin{aligned} I_1(X,Y) &= \frac{i}{2\pi Y} \exp(-i2\pi X\hat{l}) \int_{-\frac{\hat{l}\sqrt{3}}{2}}^{0} \exp(-i2\pi(X+Y\sqrt{3}/3))\,dx \\ &\quad - \frac{i}{2\pi Y} \exp(i2\pi X\hat{l}) \int_{-\frac{\hat{l}\sqrt{3}}{2}}^{0} \exp(-i2\pi(X-Y\sqrt{3}/3))\,dx \\ &= -\frac{1}{4\pi^2 Y(X+Y\sqrt{3}/3)} \exp(-i2\pi Y\hat{l})(1-\exp(i\pi(X+Y\sqrt{3}/3)\hat{l}\sqrt{3})) \\ &\quad + \frac{1}{4\pi^2 Y(X-Y\sqrt{3}/3)} \exp(i2\pi Y\hat{l})(1-\exp(i\pi(X-Y\sqrt{3}/3)\hat{l}\sqrt{3}))\,. \end{aligned} \tag{8.5}$$

Similarly, I_2 yields:

$$\begin{aligned} I_2(X,Y) &= \frac{1}{4\pi^2 Y(X+Y\sqrt{3}/3)}\exp(i2\pi Y\hat{l})(-1+\exp(-i\pi(X+Y\sqrt{3}/3)\hat{l}\sqrt{3})) \\ &\quad -\frac{1}{4\pi^2 Y(X-Y\sqrt{3}/3)}\exp(-i2\pi Y\hat{l})(-1+\exp(-i\pi(X-Y\sqrt{3}/3)\hat{l}\sqrt{3}))\,. \end{aligned} \tag{8.6}$$

Hence, the Fourier transform of the hexagonal domain is:

$$\begin{aligned} \mathcal{F}(\psi_2)(X,Y) &= I_1(X,Y)+I_2(X,Y) \\ &= -\frac{1}{2\pi^2 Y(X+Y\sqrt{3}/3)}(\cos(2\pi Y\hat{l})-\cos(\pi\hat{l}(Y-X\sqrt{3}))) \\ &\quad +\frac{1}{2\pi^2 Y(X-Y\sqrt{3}/3)}(\cos(2\pi Y\hat{l})-\cos(\pi\hat{l}(Y+X\sqrt{3})))\,. \end{aligned} \tag{8.7}$$

As explained in Chapter 4, $\mathcal{F}(\psi_2)\simeq\frac{3\sqrt{3}}{2}\hat{l}^2\exp\left(-4.8\hat{l}^2X^2-5.3\hat{l}^2Y^2)\right)$, and thus:

$$\begin{aligned} \mathcal{F}(\psi_1)\star\mathcal{F}(\psi_2) &\simeq C\frac{a\hat{l}^2}{\lambda f^2}\left(\exp(-i\pi\frac{a}{\lambda f^2}X^2)\star\exp(-4.8\hat{l}^2X^2)\right) \\ &\quad\times\left(\exp(-i\pi\frac{a}{\lambda f^2}Y^2)\star\exp(-5.3\hat{l}^2Y^2)\right) \\ &= C\ \frac{a\hat{l}^2}{\lambda f^2}\ \frac{1}{\sqrt{4.8\hat{l}^2+i\pi a/(\lambda f^2)}\sqrt{5.3\hat{l}^2+i\pi a/(\lambda f^2)}} \\ &\quad\times\exp\left(-X^2\frac{4.8\pi^2a^2\hat{l}^2+i4.8^2\lambda f^2\hat{l}^4}{\pi^2a^2+4.8^2\lambda^2f^4\hat{l}^4}\right) \\ &\quad\times\exp\left(-Y^2\frac{5.3\pi^2a^2\hat{l}^2+i5.3^2\lambda f^2\hat{l}^4}{\pi^2a^2+5.3^2\lambda^2f^4\hat{l}^4}\right). \end{aligned} \tag{8.8}$$

From here, $\mathcal{A}_{hex}$ is obtained and the rest of the derivation of the point spread function follows as presented in Chapter 4.

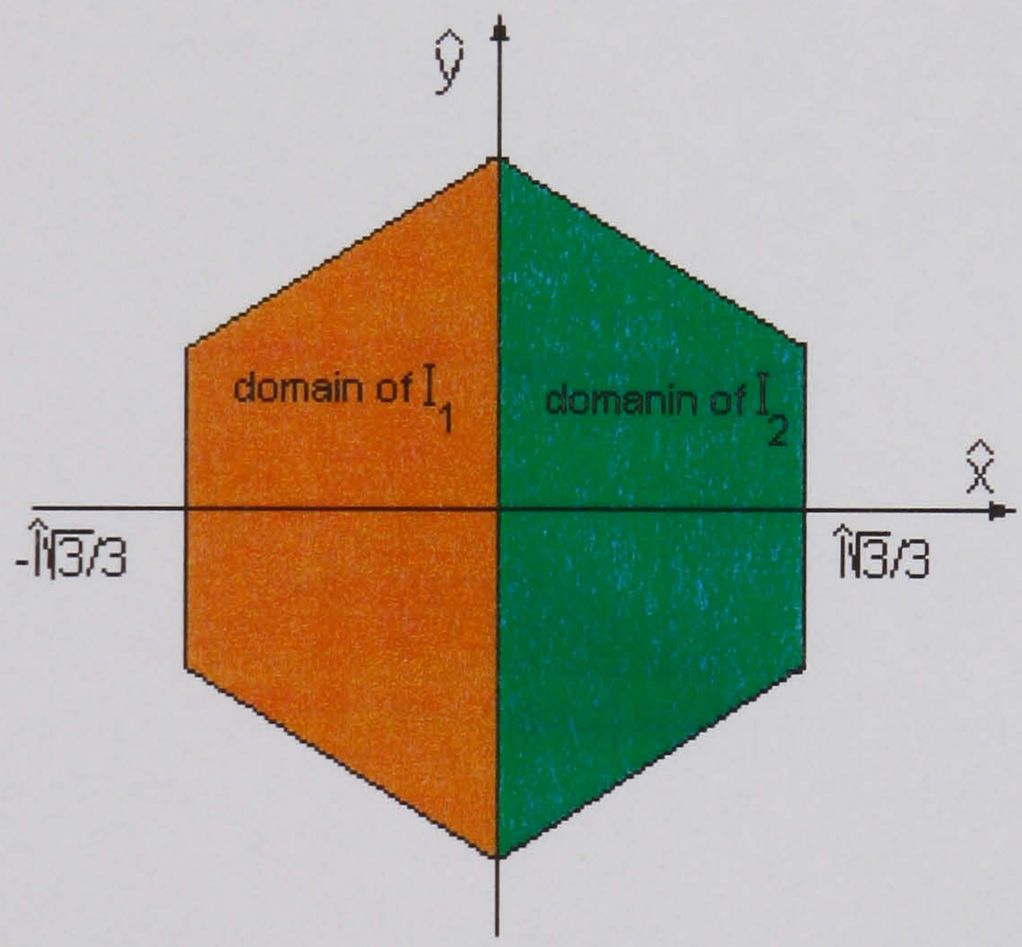
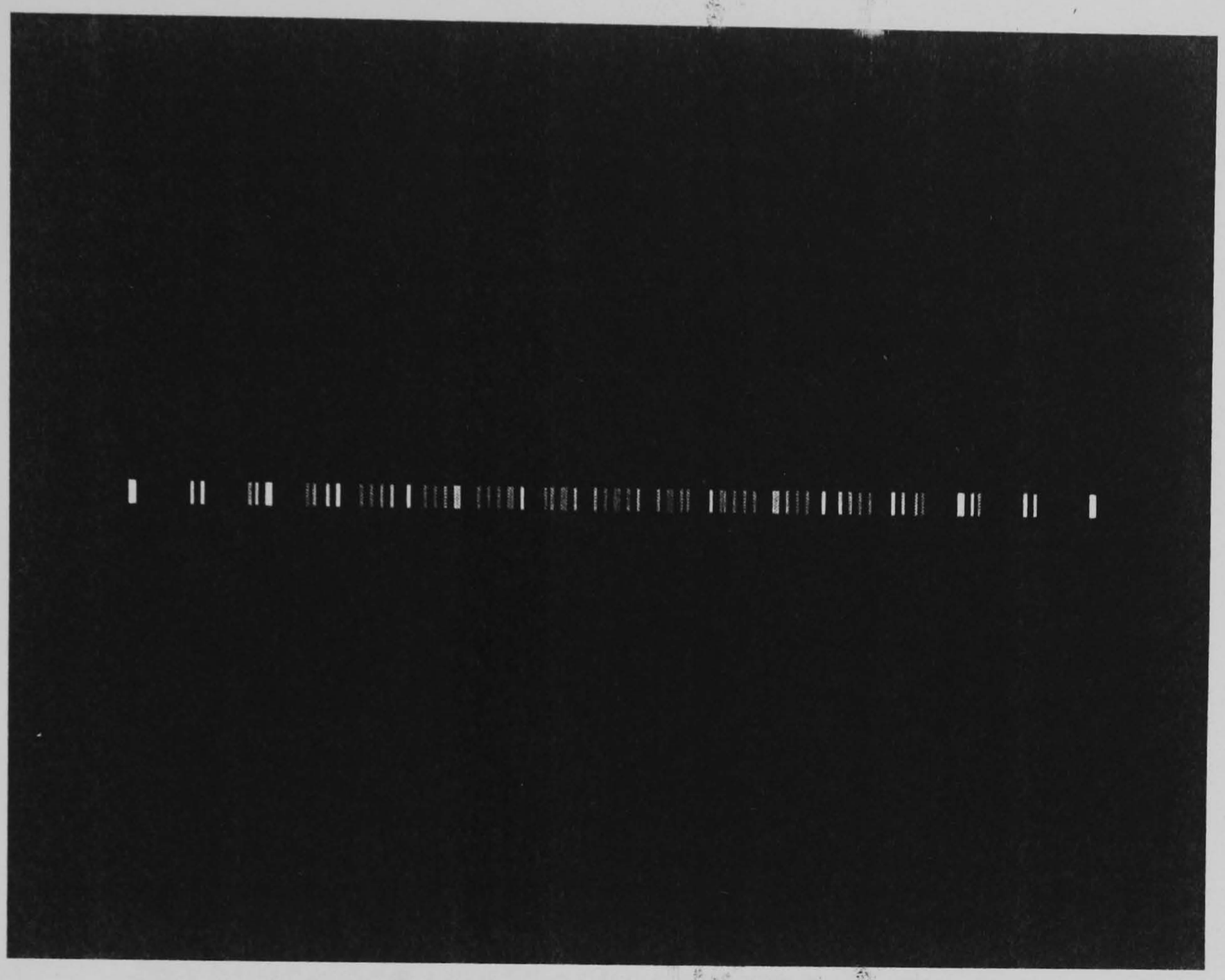
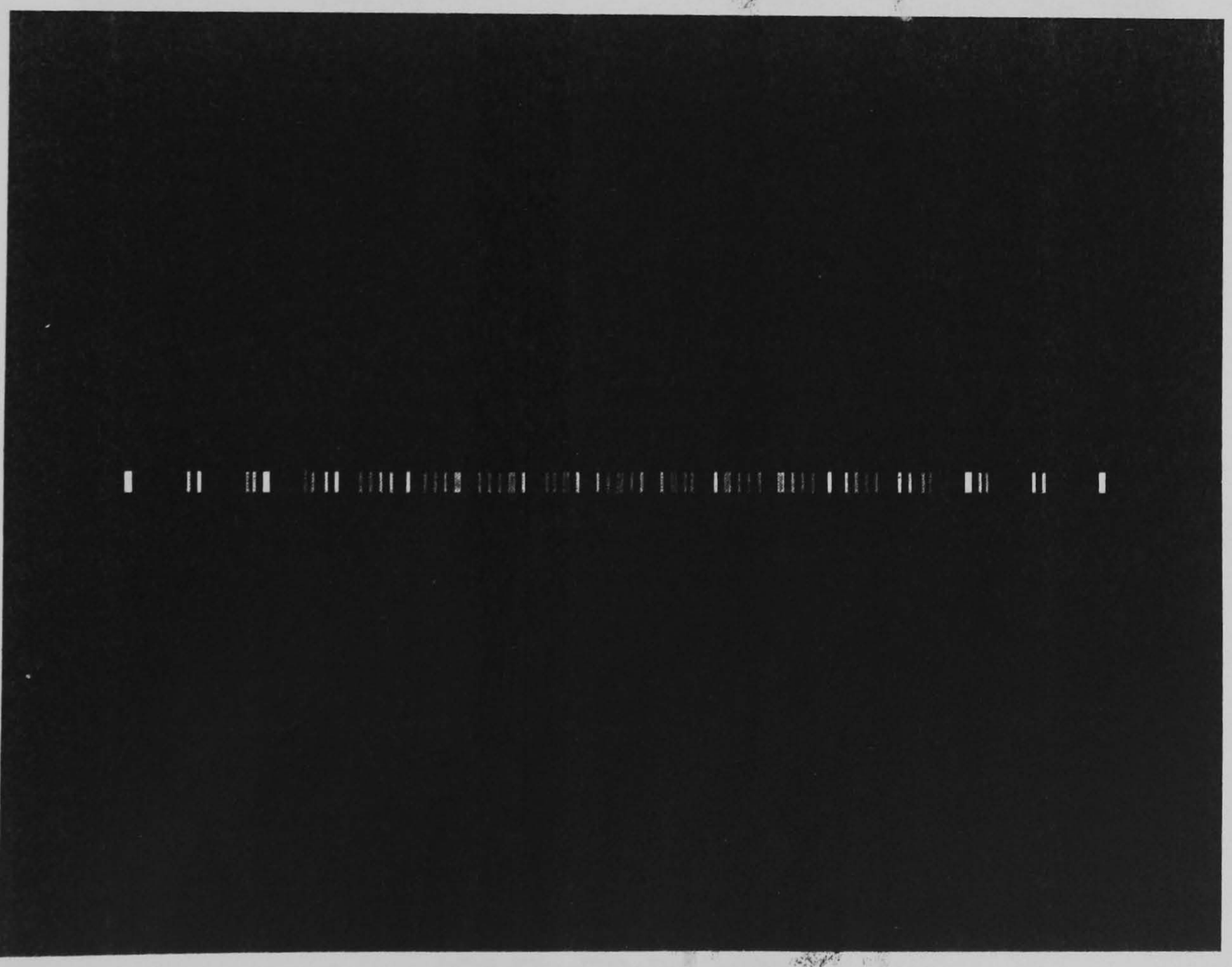


Figure 8.1: Decomposition of a hexagonal domain into two symmetric subdomains.

Appendix C: Unidirectional integral images

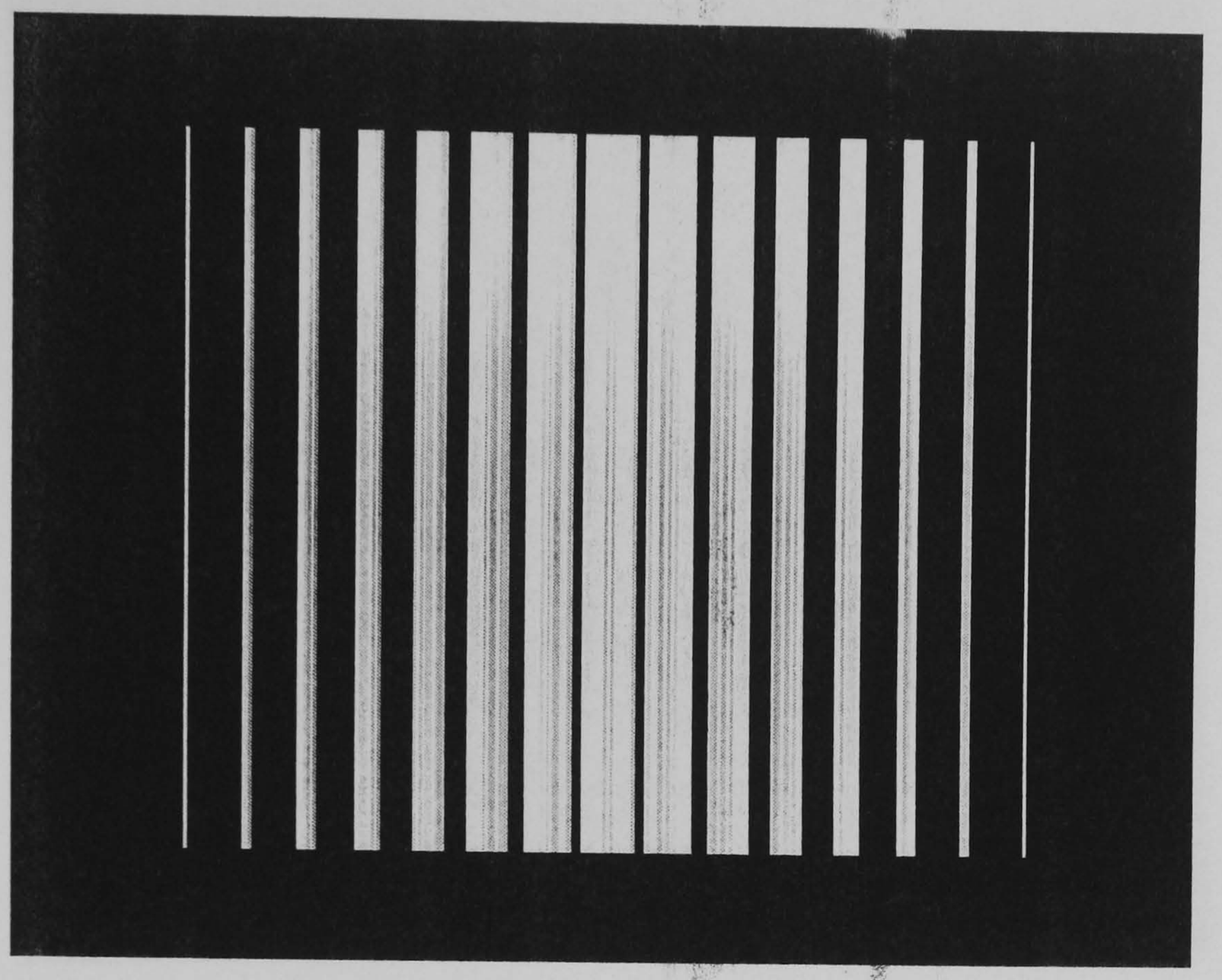


(a) Noiseless unidirectional integral image of nine luminous points

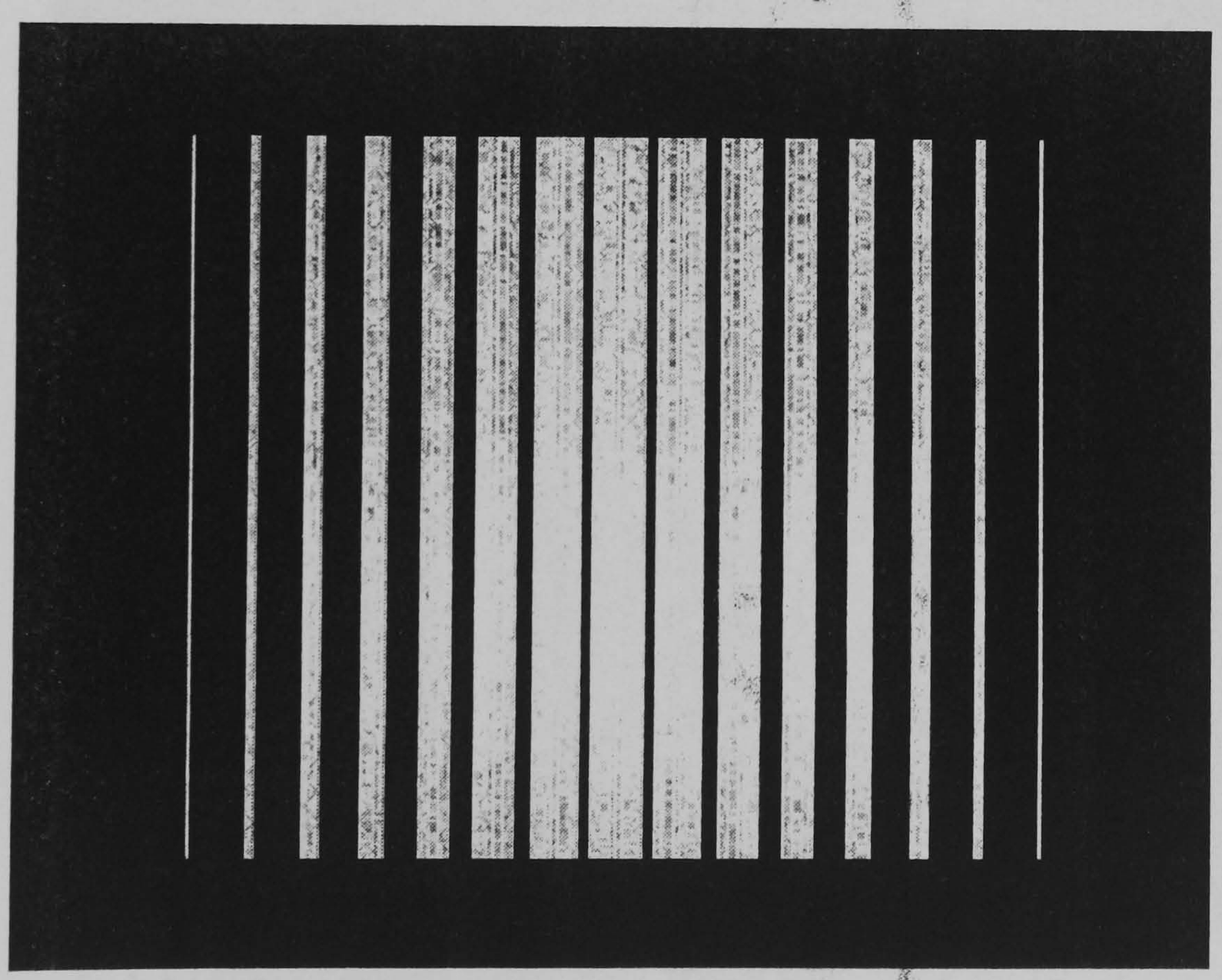


(b) Noisy unidirectional integral image of nine luminous points (SNR=20.65 dB)

Figure 8.2: (a) Noiseless and (b) noisy integral unidirectional images of nine luminous points.

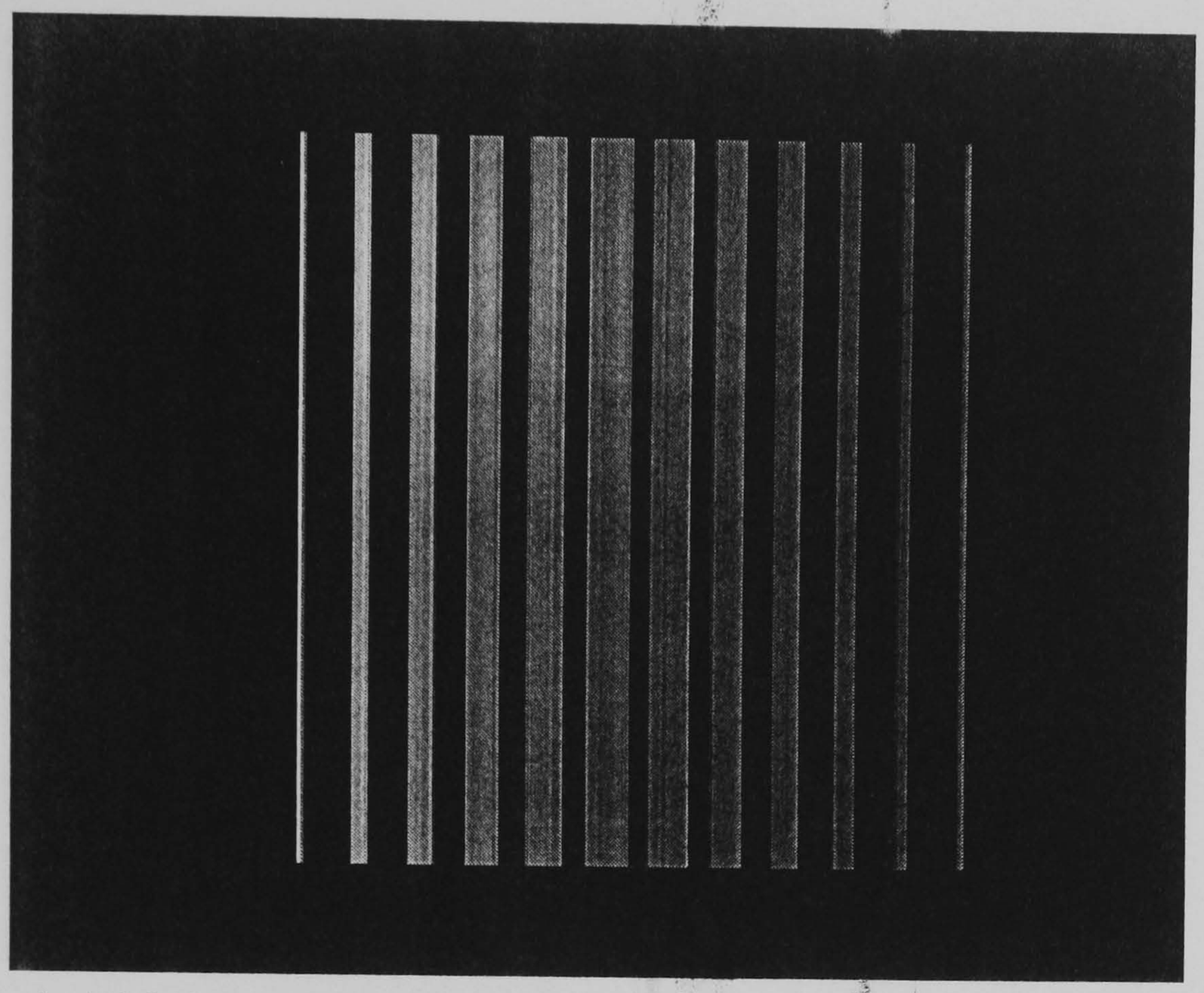


(a) Noiseless unidirectional integral image of a face parallel to the recording screen

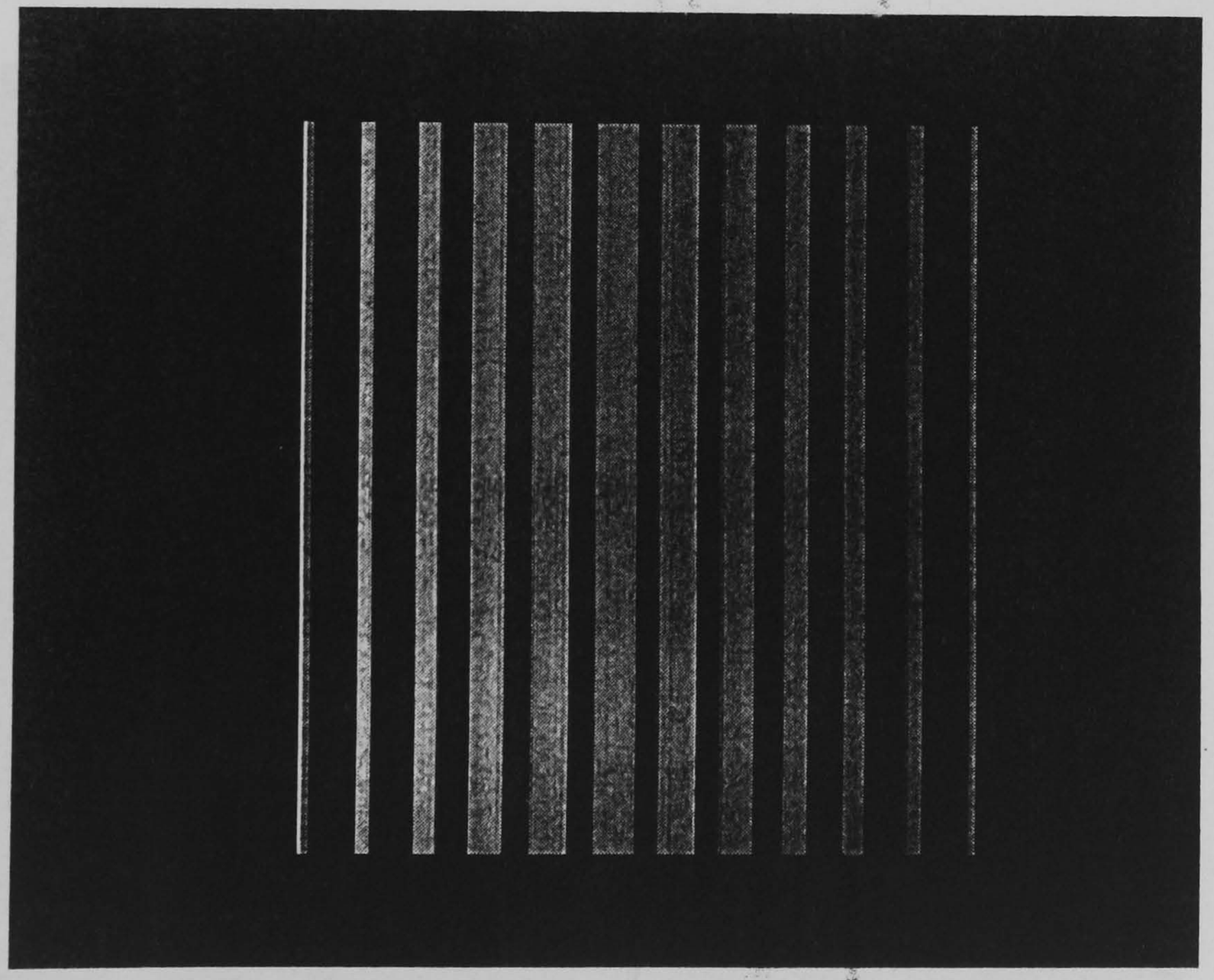


(b) Noisy unidirectional integral image of a face parallel to the recording screen (SNR=29.31 dB)

Figure 8.3: (a) Noiseless and (b) noisy unidirectional integral images of a face (plane) parallel to the recording screen.

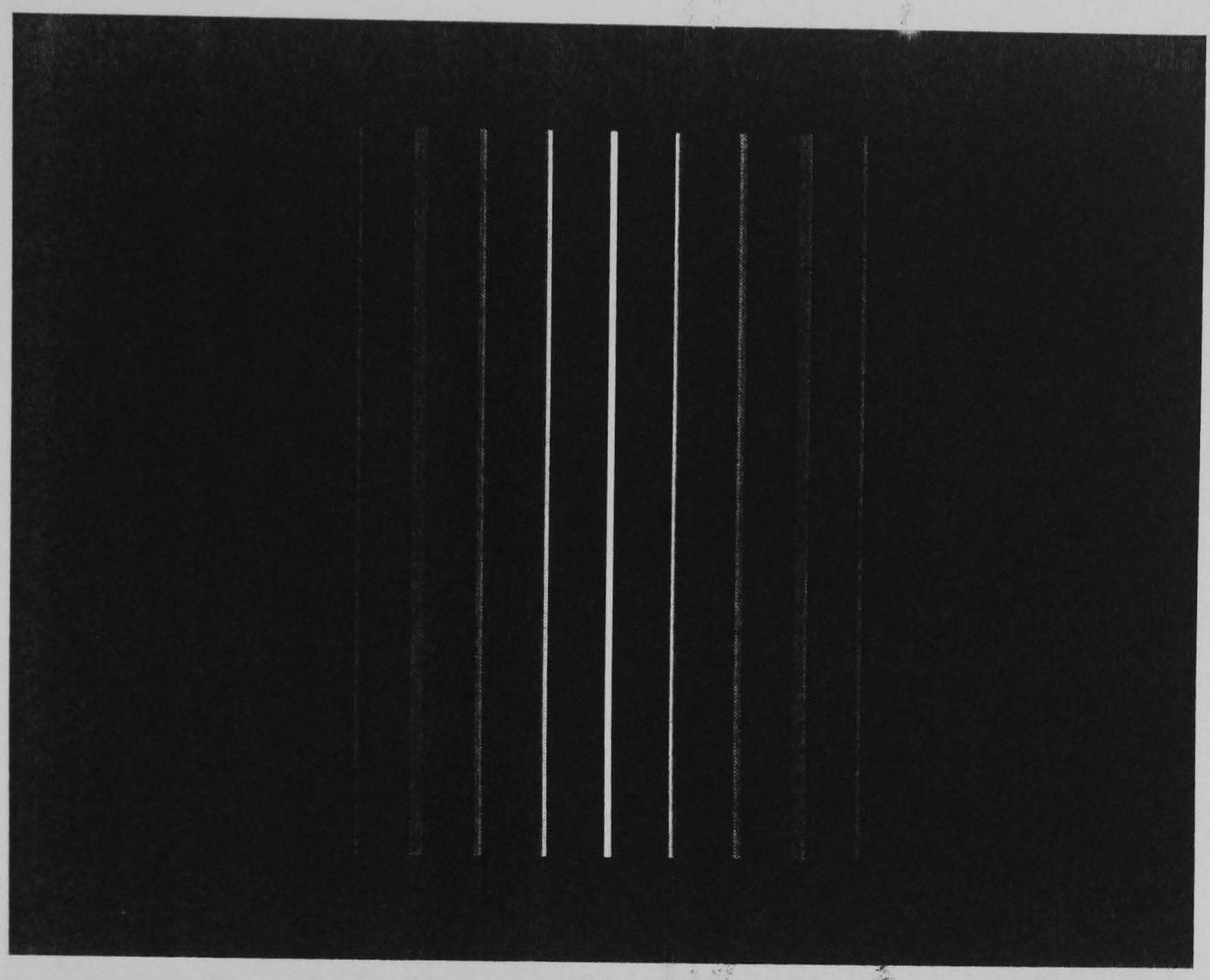


(a) Noiseless unidirectional integral image of a face inclined at 45° to the recording screen

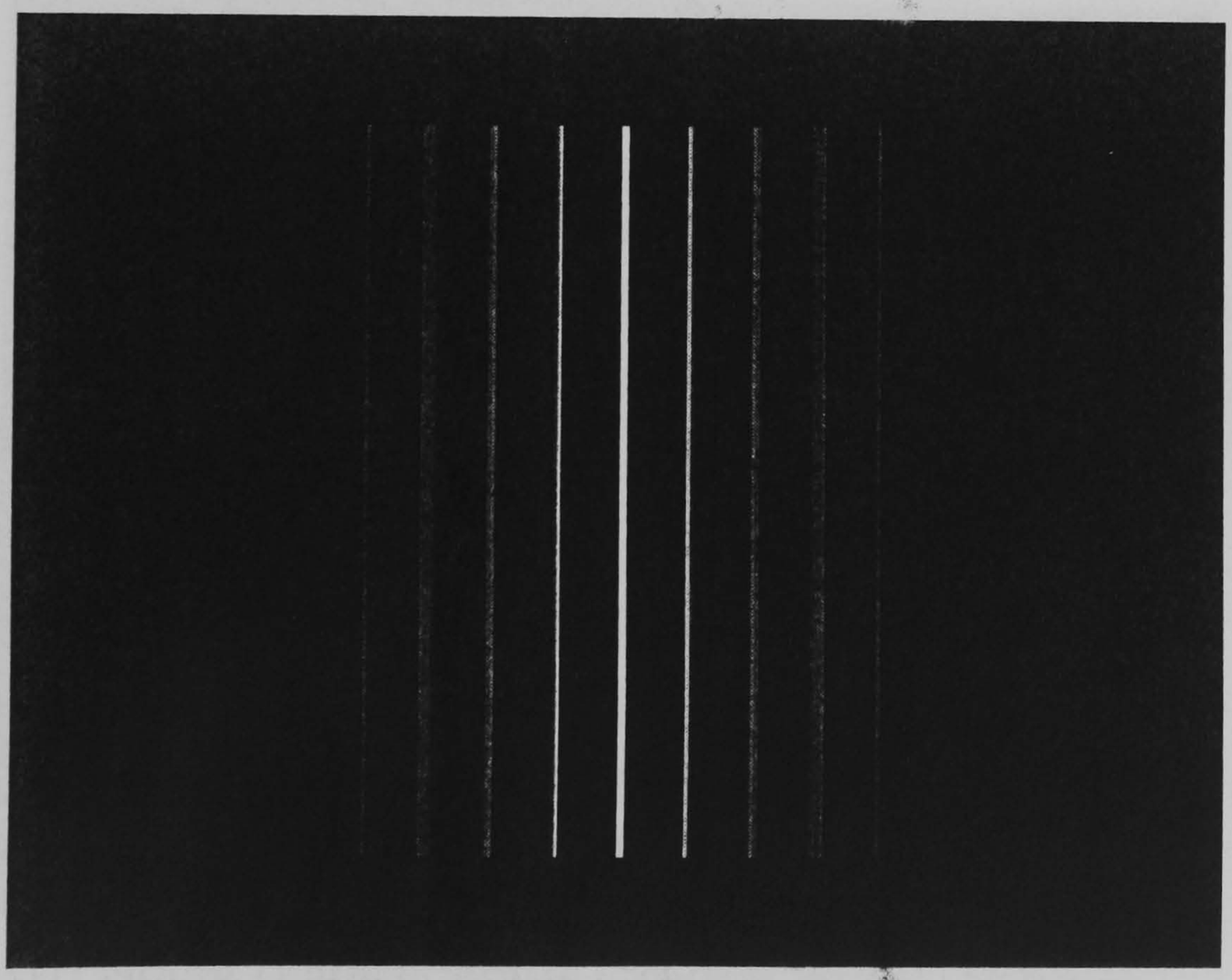


(b) Noisy unidirectional integral image of a face inclined at 45° to the recording screen (SNR=28.33 dB)

Figure 8.4: (a) Noiseless and (b) noisy unidirectional integral images of a face (plane) inclined at 45° with respect to the recording screen.

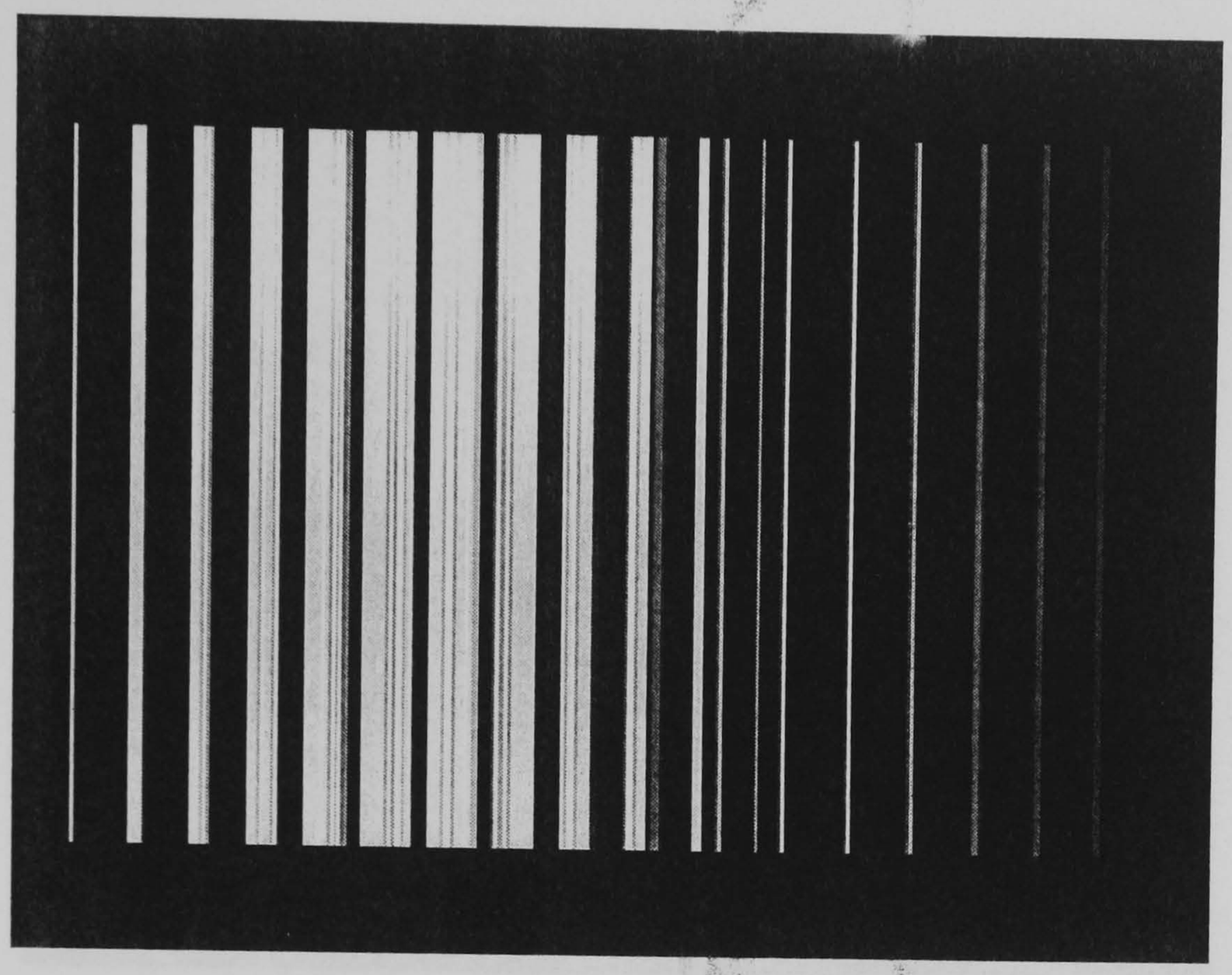


(a) Noiseless unidirectional integral image of a face orthogonal to the recording screen

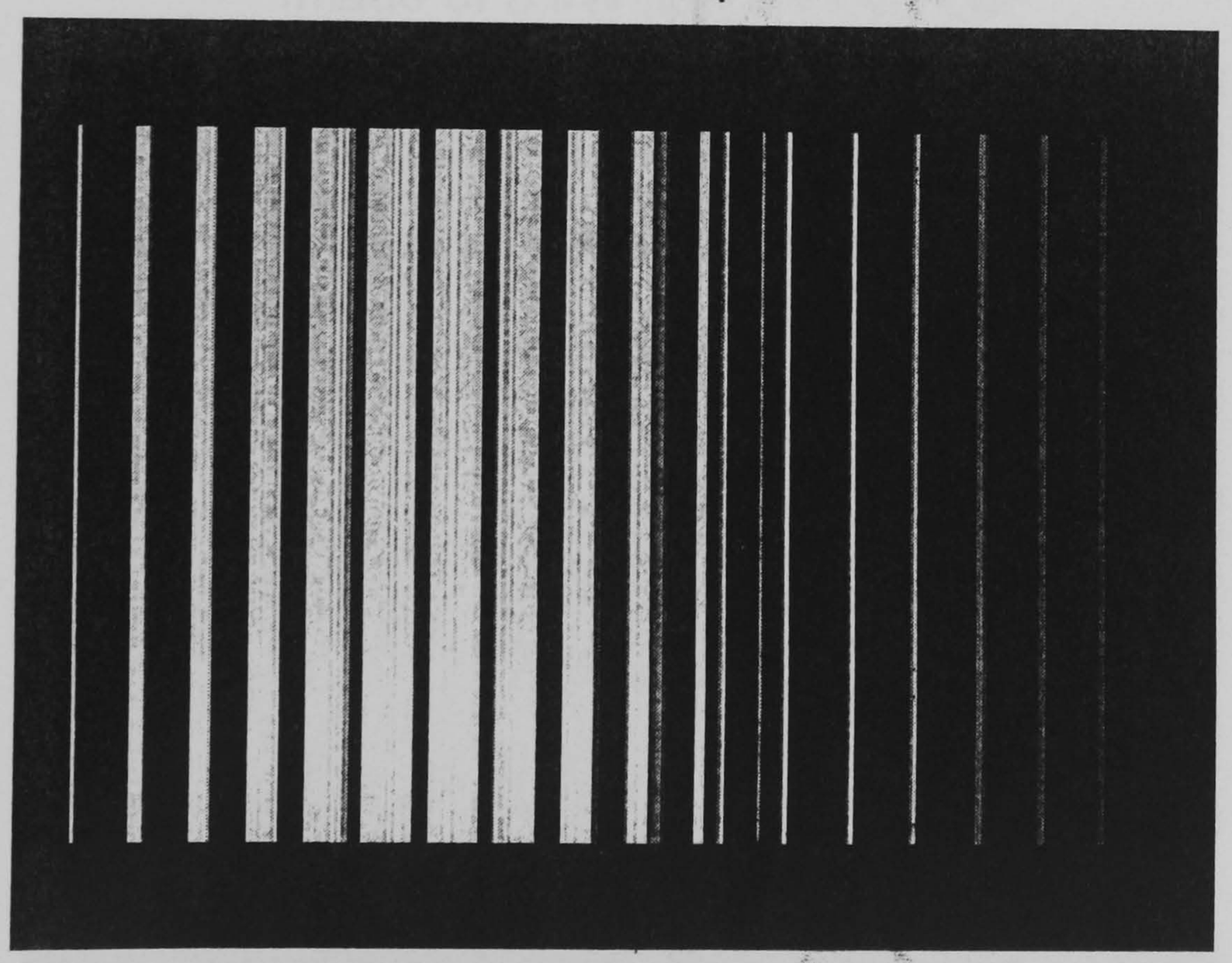


(b) Noisy unidirectional integral image of a face orthogonal to the recording screen (SNR=17.89 dB)

Figure 8.5: (a) Noiseless and (b) noisy unidirectional integral images of a face (plane) orthogonal to the recording screen.

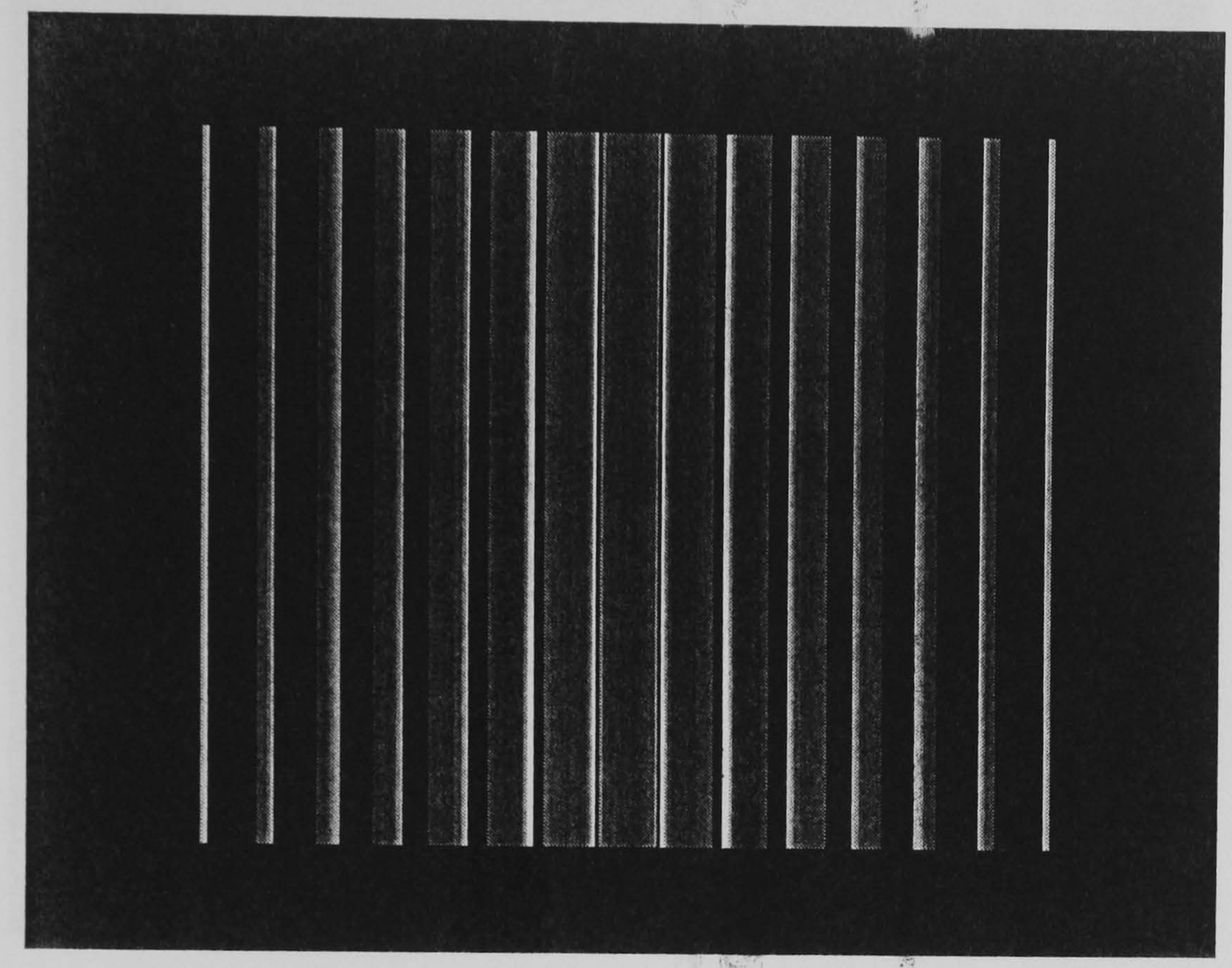


(a) Noiseless unidirectional integral image of a two transparent face object

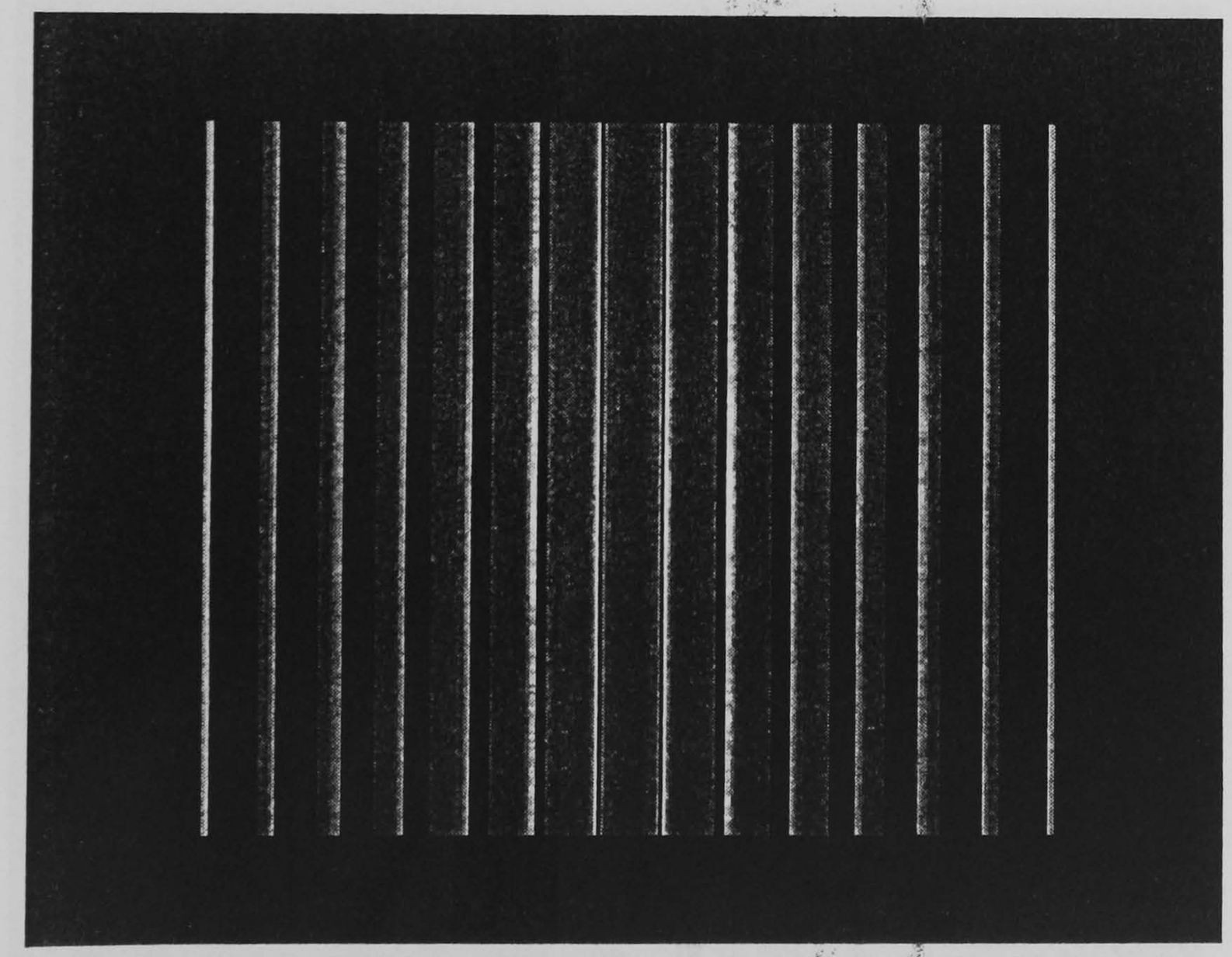


(b) Noisy unidirectional integral image of a two transparent face object (SNR=32.41 dB)

Figure 8.6: (a) Noiseless and (b) noisy unidirectional integral images of a two transparent face object space; one of the faces is parallel to the recording screen, the other is orthogonal to the recording screen.

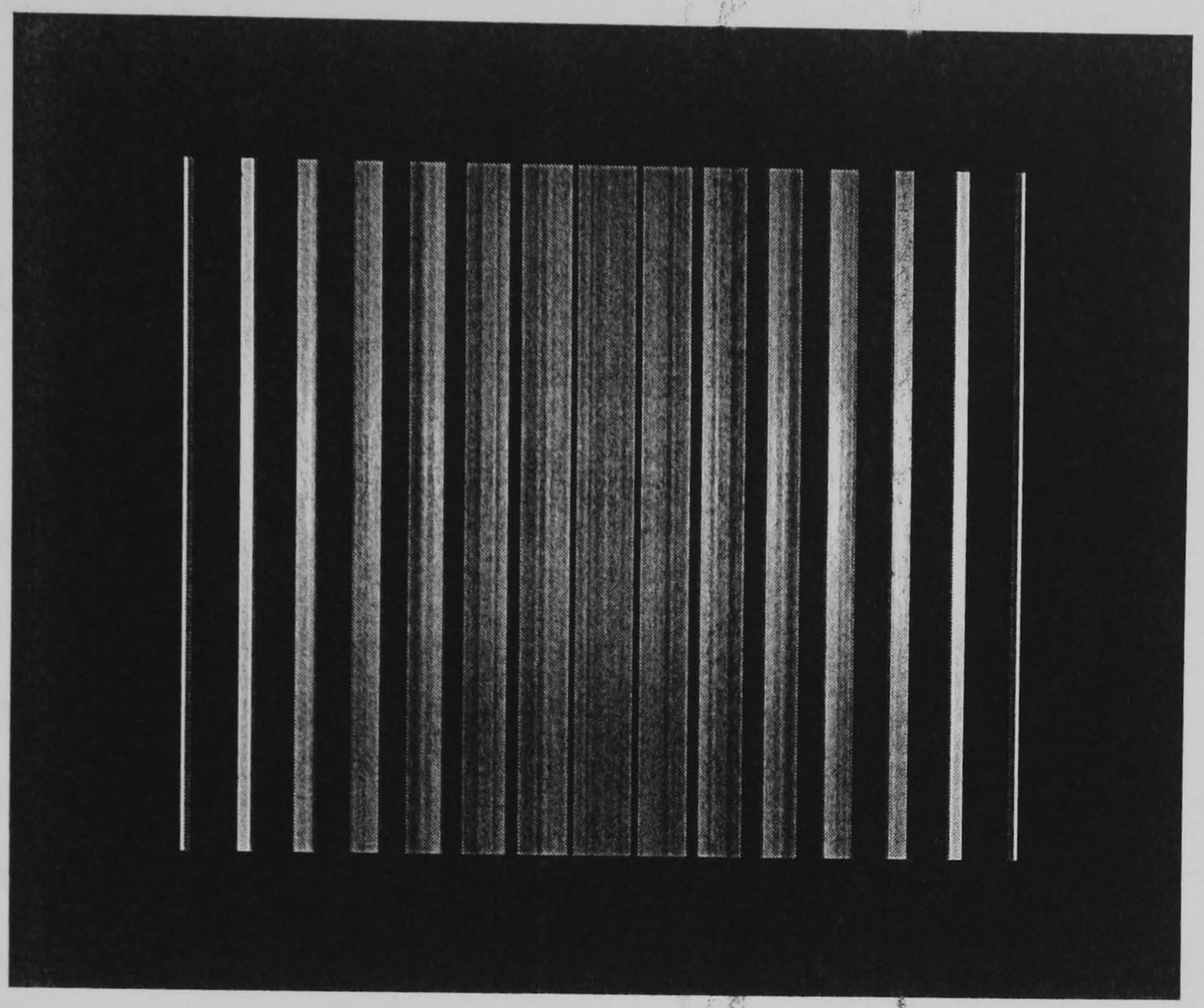


(a) Noiseless unidirectional integral image of a transparent cylinder

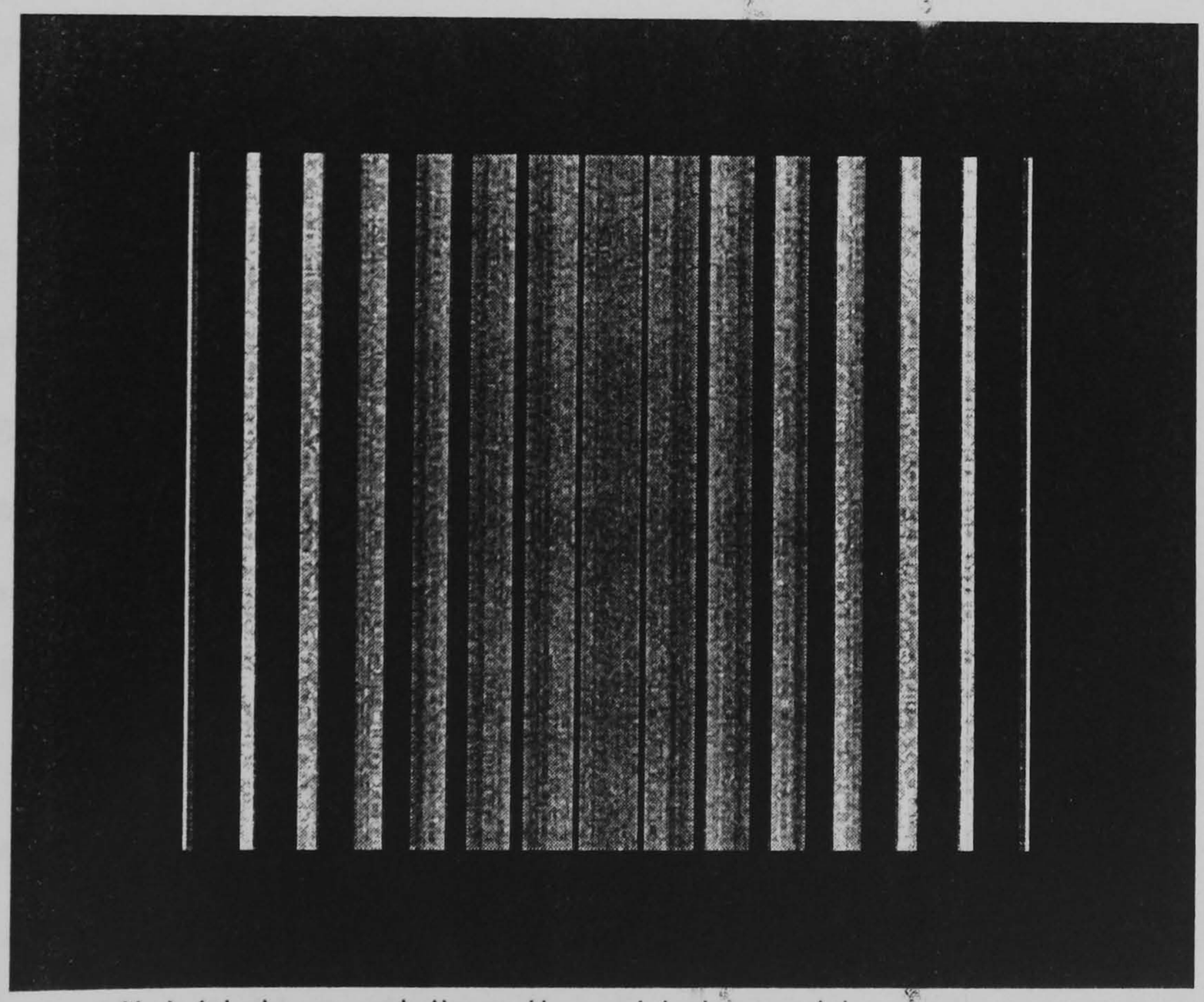


(b) Noisy unidirectional integral image of a transparent cylinder (SNR=24.60 dB)

Figure 8.7: (a) Noiseless and (b) noisy unidirectional integral images of a transparent cylinder.



(a) Noiseless unidirectional integral image of an opaque cube



(b) Noisy unidirectional integral image of an opaque cube (SNR=26.73 dB)

Figure 8.8: (a) Noiseless and (b) noisy unidirectional integral images of the visible faces of an opaque cube. The faces are inclined at $\pm 45°$ with respect to the recording screen.

List of Figures

List of Tables

Bibliography

[1] U. Amato, W. Hughes, "Maximum entropy regularization of Fredholm integral equations of the first kind", Inverse Problems, vol. 7, pp. 793-808, 1991.

[2] A.B. Bakushinskii, "A general method of constructing regularizing algorithms for a linear ill-posed equation in Hilbert space", USSR Comp. Maths. Math. Phys., vol. 7, no. 3, pp. 279-287, 1967.

[3] K.R. Barnes, "The Optical Transfer Function", Adam Hilger ltd, London, 1971.

[4] D.H. Ballard, C.M. Brown, "Computer vision", Prentice Hall, Englewood Cliffs - NJ, 1982.

[5] C. Berbente, S. Mitran, S. Zancu, "Metode numerice", Editura Tehnica, Bucuresti, 1997.

[6] M. Bertero, P. Boccacci, "Introduction to inverse problems in imaging", Bristol and Philadelphia, Institute of Physics Publishing, 1998.

[7] M. Bertero, D. Bindi, P. Boccacci, M. Cattaneo, C. Eva, V. Lanza, "Application of the projected Landweber method to the estimation of the sourse time function in seismology", Inverse Problems, vol. 13, pp. 465-486, 1997.

[8] M. Bertero, C. de Mol, G.A. Viano, "The stability of inverse problems", Inverse scattering in Optics, pp. 161-214, H.P. Baltes, ed., Springer Verlag, New York, 1980.

[9] H. Bialy, "Iterative Behandlung linearer Functionalgleichungen", Arch. Rational Mechs. Anal., vol. 4, no. 99, pp. 166-176, 1959.

[10] T.L. Boullion, P.L. Odell, "Generalised inverse matrices", Wiley-Interscience, London, 1971.

[11] M. Brewin, M. Forman, N.Davies, "Electronic capture and display of full parallax 3D images", in "Stereoscopic Displays and Virtual Reality Systems II", Proc. SPIE no. 2409, pp. 118-124, 1995.

[12] M. Brewin, "Reduced information integral imaging", PhD thesis, De Montfort University, 1998.

[13] C. Burckhardt, "Optimum parameters and resolution limitation of integral photography", J. Opt. Soc. Am., vol. 58, no. 1, pp. 71-76, 1968.

[14] C. Burckhardt, R.J. Collier, E.T. Doherty, "Formation and Inversion of pseudoscopic images", Appl. Opt., vol. 7, no. 4, pp. 627-631, 1968.

[15] A. Cauchy, "Methode generale pour la resolution des systems d'equations simultanes", C. R. Acad. Sci. Paris, vol. 25, pp. 536-538, 1847.

[16] G. Chavent, K. Kunisch, "Convergence of Tikhonov regularization for constrained ill-posed inverse problems", Inverse Problems, vol. 10, pp. 63-76, 1994.

[17] R. E. Cline, "Elements of the theory of generalized inverses for matrices", UMAP expository monograph series, Newton - Mass., 1979.

[18] P. Craven, G. Wahba, "Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation", Numer. Math., vol. 31, pp. 377-403, 1979.

[19] N. Davies, M. McCormick, Li Yang, "Three dimensional imaging systems: a new development", Appl. Opt., vol. 27, pp. 4520-4528, 1988.

[20] N. Davies, M. McCormick, "Holoscopic imaging with true 3D-content in full natural colour", J.Phot.Science, vol. 40, pp. 46-49, 1992.

[21] N. Davies, M. McCormick, M. Brewin, "Design and analysis of an image transfer system using microlens arrays", Opt. Eng., vol. 33-11, pp. 3624-3633, 1994.

[22] M. Defrise, C. de Mol, "A note on stopping rules for iterative regularization methods and filtered SVD", Advances Elect. Electron Phys. Supp., Academic Press, London, supp. 19, pp.261-268, 1987.

[23] Y.A. Dudnikov, B.K. Rozhkov, "On the possibility of using the integral photography method for stereophotogrametry of objects at close range", Sov. J. Opt. Technol., vol. 43, no. 11, pp. 646-648, 1976.

[24] Y.A. Dudnikov, B.K. Rozhkov, "Selecting the parameters of the lens array photographic system in integral photography", Sov. J. Opt. Technol., vol. 45, no. 6. pp. 349-351, 1978.

[25] Y.A. Dudnikov, B.K. Rozhkov, "The depth of sharp focusing and the resolution of a lens photographic system", Sov. J. Opt. Technol., vol. 46, no. 7, pp. 385-387, 1979.

[26] R. Durbin, D.J. Willshaw, "An analogue approach to the travelling salesman problem using an elastic net method", Nature, vol. 326, no. 16, 1987, pp. 689.

[27] P.M. Duffieux, "L'integrale de Fourier et ses applications a l'optique", New York. Wiley, 1983.

[28] B. Eicke, "Iteration methods for convexly constrained ill-posed problems in Hilbert space", Numer. Funct. Anal. and. Optimiz., vol. 13, no. 5-6, pp. 413-429, 1992.

[29] B. Eicke, A.K. Louis, R. Plato, "The instability of some gradient methods for ill-posed problems", Numer. Math., vol. 58, pp. 129-134, 1990.

[30] L. Elden, "Algorithms for the regularization of ill-conditioned least squares problems", BIT, vol. 17, pp. 134-145, 1977.

[31] H.W. Engl, K. Kunisch, A. Neubauer, "Convergence rates for Tikhonov regularization of non-linear ill-posed problems", Inverse Problems, vol. 5, pp. 63-79, 1989.

[32] D. Gabor, "A new microscope principle", Nature, no. 161, pp. 777-779, 1948.

[33] D. Gabor, "Microscopy by reconstructed wavefronts", Proc. Phys. Soc. A194, pp. 454-487, 1949.

[34] S. Gibbs, C. Arapis, C. Breiteneder, V. Lalioti, S. Mostafawy, J. Speier, "Virtual studios: an overview", IEEE MultiMedia, pp. 18-35, January - March 1998.

[35] S. Gilson, "Neural networks for computer vision", http://www.umist.ac.uk/dept/vcnrg/gilson.

[36] S.F. Gilyazov, "Iterative solution methods for inconsistent linear equations with non-selfadjoint operators", Moscow Univ. Comput. Math. Cybern., vol. 1, pp. 8-13, 1977.

[37] B. Hofmann, O. Scherzer, "Factors influencing the ill-posedness of nonlinear problems", Inverse Problems, vol. 10, pp. 1277-1297, 1994.

[38] "About holography", http://www.holophile.com/history.htm

[39] N.K. Ignat'ev, "The modes of operation of a lens array for obtaining integral photography", Sov. J. Opt. Technol., vol. 50, no. 61, pp. 6-8, 1983.

[40] H.E. Ives, "Optical properties of a Lippmann lenticulated sheet", J. Opt. Soc. Am., vol. 21, pp. 171-176, 1931.

[41] W.J. Kammerer, M.Z. Nashed, "Steepest descent for singular linear operators with nonclosed range", Applicable Analysis, vol. 1, pp. 143-159, 1971.

[42] W.J. Kammerer, M.Z. Nashed, "On the convergence of the conjugate gradient method for singular linear operator equations", SIAM J. Numer. Anal., vol. 9, no. 1, pp. 165-181, 1972.

[43] W.J. Kammerer, M.Z. Nashed, "Iterative methods for best approximate solutions of linear integral equations of the first and second kinds", J. Math, Anal. Appl., vol. 40, pp. 547-573, 1972.

[44] T. Kanade, P. Rander, P.J. Narayanan, "Virtualized reality: constructing virtual worlds from real scenes", IEEE MultiMedia, pp. 34-47, June-March 1997.

[45] M.V. Klein, T.E. Furtak, "Optics" (John Wiley&Sons, New York, 1986).

[46] S.Y. Kung, "Digital neural networks", Prentice-Hall, 1993.

[47] K. Kunisch, J. Zou, "Iterative choices of regularization parameters in linear inverse problems", Inverse Problems, vol. 14, pp. 1247-1264, 1998.

[48] L. Landweber, "An iteration formula for Fredholm integral equations of the first kind", Amer. J. Math., vol. 73, pp. 615-624, 1951.

[49] E.N. Leith, J. Upatnieks, "Reconstructed wavefronts and communication theory", J. Opt. Soc. Amer., vol. 52, no. 10, pp. 1123-1130, 1962.

[50] E.N. Leith, J. Upatnieks, "Wavefront reconstruction with continuous-tone objects", J. Opt. Soc. Amer., vol. 53, no. 12, pp. 1377-1381, 1963.

[51] E.N. Leith, J. Upatnieks, "Wavefront reconstruction with diffused illumination and three-dimensional objects", J. Opt. Soc. Amer., vol. 54, no. 11, pp. 1295-1301, 1964.

[52] P. Linz, "A new numerical method for ill-posed problems", Inverse problems, vol. 10, pp. L1-L6, 1994.

[53] G. Lippmann, "La Photographie integrale", Comtes Rendus, Academie des Sciences, vol.146, 1908, pp. 446-451.

[54] A.K. Louis, "Approximate inverse for linear and some nonlinear problems", Inverse Problems, vol. 12, pp. 175-190, 1996.

[55] D.G. Luenberger, "Linear and nonlinear programming", Addison-Wesley Publ., London, 1984.

[56] S. Manolache, A. Aggoun, M. McCormick, N. Davies, "A mathematical model of a 3D-lenticular integral recording system", Proc. IEEE Vision, Modeling and Visualization, Erlangen, pp. 51-58, 1999.

[57] S. Manolache, A. Aggoun, M. McCormick, N. Davies, S.Y. Kung, "Analytical model of a three-dimensional integral image recording system that uses circular- and hexagonal-based spherical surface microlenses", J. Opt. Soc. Am. A, vol. 18, no. 8, pp. 1814-1821, 2001.

[58] S. Manolache, M. McCormick, S.Y. Kung, "Hierarchical adaptive regularization method for depth extraction from planar recdording of 3D-integral images", Proc. IEEE Int. Conf. Acoust. Speech Sign. Proc. (ICASSP) (CD-ROM), vol. 3, pp. III-1433-III-1436, Salt Lake City, 2001.

[59] S. Manolache, S.Y. Kung, M. McCormick, A. Aggoun, "3D-object space reconstruction from planar recorded data of 3D-integral images", to appear in J. VLSI Signal Proc. Systems, Kluwer Acad. Publ., 2001.

[60] S. Manolache, S.Y. Kung, M. McCormick, A. Aggoun, "Regularization methods for object space reconstruction from 3D-integral images", Actes des Conf. Numerisation 3D, paper 12, Paris, 2001.

[61] J. Marsden, "Calculus", Springer Verlag, New York, 1984.

[62] David Marshall, "3D Imaging", http://www.cs.cf.ac.uk/Dave/Visionlecture.

[63] L. Matthies, M. Okutomi, "A Bayesian foundation for active stereo vision", Proc. SPIE, vol. 1198 Sensor Fusion II: Human and Machine Strategies, pp. 62-74, 1989.

[64] M. McCormick, N. Davies, "3D-worlds", Physics World, pp. 42-46, June 1992.

[65] K. Miller, "Least squares methods for ill-posed problems with a prescribed bound", SIAM J. Math. Anal., vol. 1, no. 1, pp. 52-74, 1970.

[66] S. Moezzi, L. Tai, P. Gerard, "Virtual view generation for 3D digital video", IEEE MultiMedia, pp. 19-26, June-March 1997.

[67] V.A. Morozov, "Methods for solving incorrectly posed problems", Springer Verlag, New York, 1984.

[68] P.J. Narayanan, P. Rander, T. Kanade, "Synchronous capture of image sequences from multiple cameras", CMU RI TR 95-25, Dec. 1995 (www.cs.cmu.edu/afs/cs/project/VirtualizedR/www/papers.html).

[69] A. Neubauer, "Tikhonov-regularization of ill-posed linear operator equations on closed convex sets", J. Approx. Theory, vol. 53, pp. 304-320, 1988.

[70] A. Neubauer, O. Scherzer, "Regularization for curve representations: uniform convergence for discontinuous solutions of ill-posed problems", SIAM J. Appl. Math., vol. 58, no. 6, pp. 1891-1900, 1998.

[71] A. Neumaier, "Solving ill-conditioned and singular linear systems: a tutorial on regularization", SIAM Rev., vol. 40, no. 3, pp. 636-666, 1998.

[72] Y. Ohta, T. Kanade, "Stereo by intra- and inter-scanline search using dynamic programming", IEEE Trans. Pattern Anal. Mach. Intel., vol. PAMI-7, no. 2, pp. 139-154. 1985.

[73] F. Okano, J. Arai, H. Hoshino, I. Yuyama, "Three-dimensional video system based on integral photography", Opt. Eng., vol. 38, no. 6, pp. 1072-1077, 1999.

[74] T. Okoshi, "Three dimensional imaging techniques", Academic Press, London, 1976.

[75] T. Okoshi, "Three dimensional displays", Proc. IEEE, vol. 68, no. 5,pp. 548-564, 1980.

[76] M. Okutomi. T. Kanade, "A multiple baseline stereo", IEEE Trans. Pattern Anal. Mach. Intel., vol. 15. no. 4. pp. 353-363. 1993.

[77] J.M. Ortega, W.C. Rheinboldt, "Iterative solutions of nonlinear equations in several variables", Academic Press, 1970.

[78] C. Outwater, V. Hamersveld, "Practical Holography", Dimensional Arts Inc. 1995-99, http://www.holo.com/holo/book/book1.html.

[79] M. Piana, M. Bertero, "Projected Landweber method and preconditioning", Inverse Problems, vol. 13, pp. 441-463, 1997.

[80] J.W. Rayleigh, "On the theory of optical images, with special reference to microscope", Phil. Mag., vol. 42, pp. 167-195, 1896 (appud "Scientific Papers", vol. IV, pp. 235-260, CUP, Cambridge).

[81] B.K. Rozhkov, "The use of transformation properties of the lens-array system in integral photography", Sov. J. Opt. Technol., vol. 54, no. 2, pp. 67-70, 1987.

[82] J.L.C. Sanz, T.S. Huang, "Unified Hilbert space approach to iterative least-squars linear signal restoration", J. Opt. Soc. Am., vol. 73, no. 11, pp. 1455-1465, 1983.

[83] Gh. Siretchi, "Calcul diferential si integral", Editura Stiintifica si Enciclopedica, Bucuresti, 1985.

[84] "The remote sensing core curricullum", http://www.research.umbc.edu/tbenja1.

[85] E.G. Steward, "Fourier optics - an introduction", Ellis Horwood ltd., 1987.

[86] O.N. Strand, "Theory and methods related to the singular-function expansion and Landweber's iteration for integral equations of the first kind", SIAM J. Numer. Anal., vol. 11, no. 4, pp. 798-825, 1974.

[87] U. Tautenhahn, "Error estimates for regularized solutions of non-linear ill-posed problems", Inverse Problems, vol. 10, pp. 485-500, 1994.

[88] A.N. Tikhonov, V.Y. Arsenin, "Solutions of ill-posed problems", Wiley, New York, 1977.

[89] A.N. Tikhonov, "On the regularization of ill-posed problems", Dokl. Akad. Nauk. SSSR, vol. 153, no. 1, pp. 49-52, 1963.

[90] A.N. Tikhonov, "Solution of incorrectly formulated problems and the regularization method", Soviet Math. Dokl., vol. 4, 1963, pp. 1035-1038, 1963.

[91] B.L. Tseng, D. Anastassiou, "Multi-viewpointvideo coding with MPEG-2 compatibility", IEEE Trans. Circuits and Systems for Video Tech., vol. 6, no. 4, pp. 414-419, 1996.

[92] G.M. Vainikko, "The critical level of discrepancy in regularization methods", USSR Comput. Maths. Math. Phys., vol. 23, no. 6, pp. 1-9, 1983.

[93] N.A. Valyus, "Stereoscopy", Focal Press, London, 1966.

[94] V.V. Vasin, "Iterative methods for solving ill-posed problems with a priori information in Hilbert spaces", USSR Comput. Maths. Math. Phys., vol. 28, no. 4, pp. 6-13, 1988.

[95] G.D. de Villiers, B. McNally, E.R. Pike, "Positive solutions to linear inverse problems", Inverse Problems, vol. 15, pp. 615-635, 1999.

[96] G. Wahba, "Practical approximate solutions to linear operator equations when the data are noisy", SIAM J. Numer. Anal., vol. 14, no. 4, pp. 651-667, 1977.

[97] Y. Wang, T. Xiao, "Fast realization algorithms for determining regularization parameters in linear inverse problems", Inverse Problems, vol. 17, pp. 281-191, 2001.

Publications in full text ([56]-[60])

The following publications are presented in full text:

57 S. Manolache, A. Aggoun, M. McCormick, N. Davies, S.Y. Kung, **Analytical model of a three-dimensional integral image recording system that uses circular and hexagonal based spherical surface microlenses**, Journal of the Optical Society of America A, vol. 18, no. 8, pp. 1814-1821, 2001.

59 S. Manolache, S.Y. Kung, M. McCormick, A. Aggoun, **3D-object space reconstruction from planar recorded data of 3D-integral images**, to appear in Journal of VLSI Signal Processing Systems, Kluwer Acad. Publ., 2001.

56 S. Manolache, A. Aggoun, M. McCormick, N. Davies, **A mathematical model of a 3D-lenticular integral recording system**, Proceedings of IEEE Vision, Modeling and Visualization Conference, Erlangen, pp. 51-58, Nov. 1999.

58 S. Manolache, M. McCormick, S.Y. Kung, **Hierarchical adaptive regularization method for depth extraction from planar recdording of 3D-integral images**, Proceedings of IEEE International Conference for Acoustics, Speech and Signal Processing (ICASSP), vol. 3, pp. III-1433-III-1436, May 2001.

60 S. Manolache, S.Y. Kung, M. McCormick, A. Aggoun, **Regularization methods for object space reconstruction from 3D-integral images**, Actes des Conferences Numerisation 3D, paper 12, Paris, April 2001.

Analytical model of a three-dimensional integral image recording system that uses circular- and hexagonal-based spherical surface microlenses

Silvia Manolache, Amar Aggoun, Malcolm McCormick, and Neil Davies

Department of Electrical and Electronic Engineering, De Montfort University, The Gateway, Leicester LE1 9BH, UK

S. Y. Kung

Department of Electrical Engineering, Princeton University, Room B230, Engineering Quadrangle, Princeton, New Jersey 08544

Received August 1, 2000; revised manuscript received January 11, 2001; accepted January 16, 2001

A mathematical model for a three-dimensional omnidirectional integral recording camera system that uses either circular- or hexagonal-based spherical surface microlens arrays is derived. The geometry of the image formation and recording process is fully described. Matlab is then used to establish the number of recorded micro-intensity distributions representing a single object point and their dependence on spatial position. The point-spread function for the entire optical process for both close and remote imaging is obtained, and the influence of depth on the point-spread dimensions for each type of microlens and imaging condition is discussed. Comparisons of the two arrangements are made, based on the illustrative numerical results presented.

OCIS codes: 000.4430, 080.2740, 110.2990, 110.4850, 120.4820.

1. INTRODUCTION

Three-dimensional (3D) imaging systems employing incoherent light sources have many advantages, not the least being application in a nonspecialized environment. Integral photography (IP), which is based on using microlens arrays for capturing images, is an important method available in this area and was first proposed by Lippman[1] in 1908. It creates true-color 3D images that do not require special viewing spectacles, displays continuous parallax, and does not exhibit flipping. The method was further investigated and improved by Ives, who re-exposed to overcome the pseudoscopy of the image produced by Lippmann's system.[2] Consequently, to produce an orthoscopic image in replay, a two-stage process is required. Burchhardt and co-workers[3,4] proposed an analytical approach to determining the optimum parameters of an IP system. Okoshi[5] also studied the optics of IP in the particular case of unidirectional IP, which uses an array of cylindrical surface microlenses as a direction-selective screen. Both Burkhardt and Okoshi confined their analysis to considering the effect of only one microlens of the array. Recently, there have been vigorous efforts to develop an integral system suitable for television applications,[6,7] and pick-up and display systems have been reported.[7,8]

The optics of an advanced form of integral imaging system—that employs a two-tier optical network—in which a true 3D optical reconstruction of a scene is transferred to the capture device has been described in detail by Davies and McCormick and co-workers.[6,9,10] The applications of this system are multiple, ranging from 3D television to robotic vision and medical displays, so that the discovery of the rules relating the system parameters to the formation and quality of the 3D image is a crucial problem. To our knowledge, to date a comprehensive analysis of the two-tier camera system has not been reported and the variation of the recorded intensity distribution with respect to the point depth has not been presented.

In this paper we seek to rectify the situation by deriving a mathematical model of the omnidirectional 3D integral recording camera system using either circular-based or hexagonal-based spherical surface microlens arrays in the autocollimating transmission screen and encoding screen. The close relationship between the image morphology and the 3D recorded scene composition is analyzed, and comparative considerations for the two variants of the system are presented.

2. TWO-TIER OPTICAL NETWORK

The optical arrangement, shown in Fig. 1, comprises two macrolens arrays (MA1 and MA2) placed equidistantly behind and in front of an autocollimating transmission screen (ATS). The ATS is made up of two microlens arrays separated by their joint focal distance. The recording plane, a photographic plate whose position coincides with the focal plane of an additional microlens array (RA), lies within the optical model.

Within this paper, a few terms are used to express the type of information distribution associated with an imaged scene: an image is defined to be a point-by-point reconstruction of the object space; an optical model is considered to be formed by intersecting modulated ray

0740-3232/2001/081814-08$15.00

bundles, i.e., an integral image; a voxel represents a resolvable volumetric point in space; the final recorded data are a sampled set of disparate lenslet encoded microintensity distributions.

The object is imaged by the input macrolenses that form small images that occur at or near the central double-microlens screen (the ATS). The screen inverts each intermediary image, which is then transmitted to the corresponding output macrolens to form an axially inverted optical model. The output macrolenses image the optical models to their correct spatial location. The final integrated image, formed by superposed optical models projected by the second macrolens array, is a true 3D optical 1:1 reconstruction of the object. This can be recorded as a sampled data set possessing either horizontal parallax (if cylindrical surface microlenses have been used in the recording array) or continuous parallax (in the case of the circular, hexagonal, or square-based microlens arrangement).

The integral image is recorded at the focal plane of a recording microlens array as a lenslet encoded spatial distribution that contains omnidirectional information relating to the object space. Each microlens of the recording array samples a fractional part of the scene; consequently, many microlenses record directional information of the scene from different angles of view. Therefore parallax information about any particular point is distributed over the recording plane. Redisplay of the full spatial model as a real 3D image can be effected by overlaying the sampled data set by an integral decoding element that is identical to the original recording array and reversing the illumination.

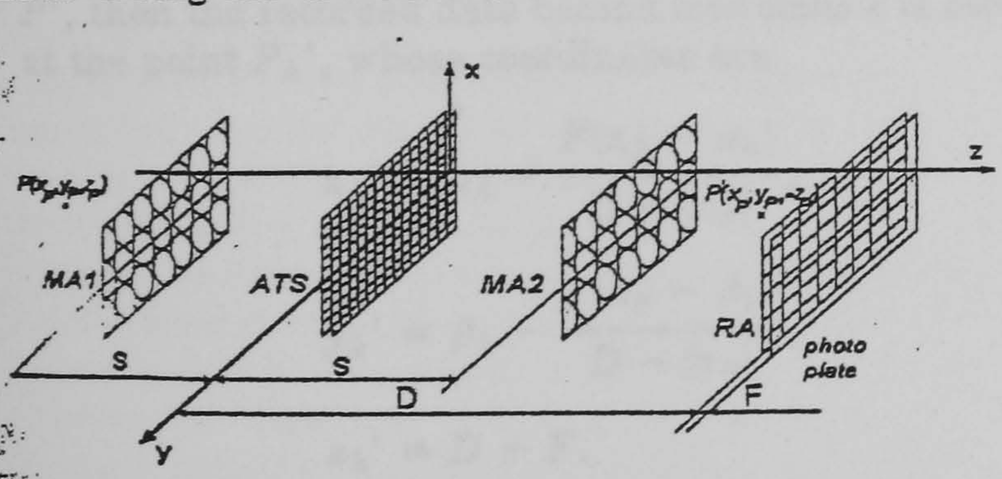


Fig. 1. Two-tier optical system.

3. GEOMETRY OF INTEGRAL IMAGE FORMATION AND RECORDING

All computations and formulas are related to the Cartesian coordinate system *Oxyz* depicted in Fig. 1. The z axis denotes the depth direction, and the x, y axes describe the lateral positions. The macrolenses and microlenses are indexed in the sense indicated by the positive axis *Oy*.

Any imaged point $P(x_p, y_p, z_p)$ has a number of states before the recording stage. The macrolneses of the first array present a first set of images of P near the autocollimating transmission screen. These images are located at the points $P_k{}^1$ having the coordinates

$$x_k{}^1 = \frac{a_k(z_P + S) + x_P f_1}{f_1 + S + z_P},$$

$$y_k{}^1 = \frac{b_k(z_P + S) + y_P f_1}{f_1 + S + z_P},$$

$$z_k{}^1 = \frac{z_P f_1 - z_P S - S^2}{f_1 + S + z_P}, \tag{1}$$

for any macrolens k in the first array, where f_1 is the focal length of the macrolenses, (a_k, b_k, S) denote the coordinates of macrolens k's center, and S is the distance between a macrolens array and the central plane *Oxy* [see Figs. 1 and 2(a)].

These intermediary images become object points for the ATS. The light emerging from them is refracted and diffracted through the microlens apertures of the first array, producing an intensity distribution in the common focal plane of the two microlens arrays. (The area between microlenses is opaque.) This intensity distribution is then transmitted by the second microlens array through a symmetric optical process. As a result, the ATS operates by refracting the transmitted ray at an equal angle of opposite sign to the angle of incidence of the impinging ray. In this way, every "object" point $P_k{}^1$ in front of the screen has a quasi-equidistant voxel representation centered at the point $P_k{}^2(x_k{}^2, y_k{}^2, z_k{}^2)$ behind the screen, where

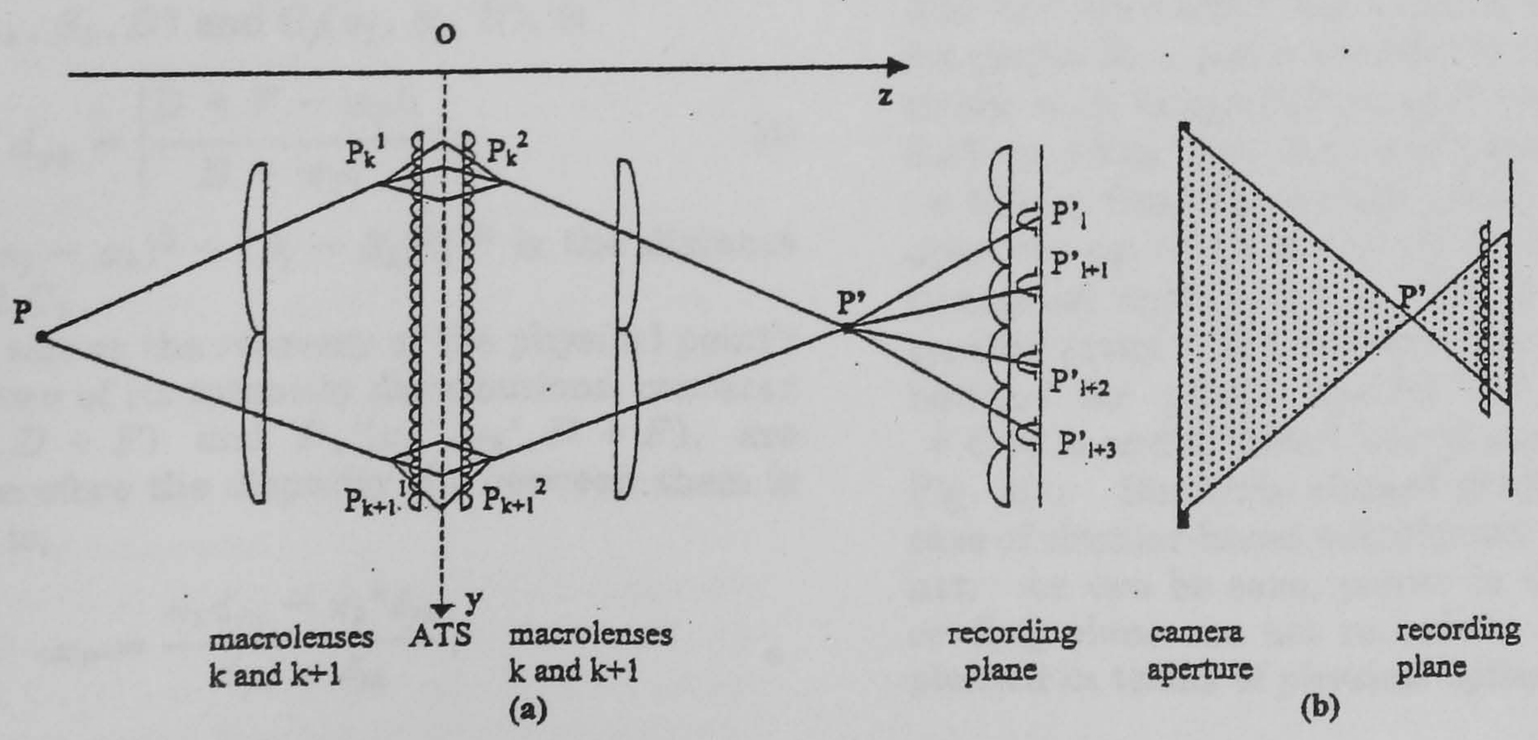


Fig. 2. Transverse sections of the camera arrangement with a plane $x = x_P$: (a) schematic representation of the intermediary, [illegible] and recorded optical models of a physical point P; (b) microlenses of the recording array that record a certain optical recon[illegible] P'.

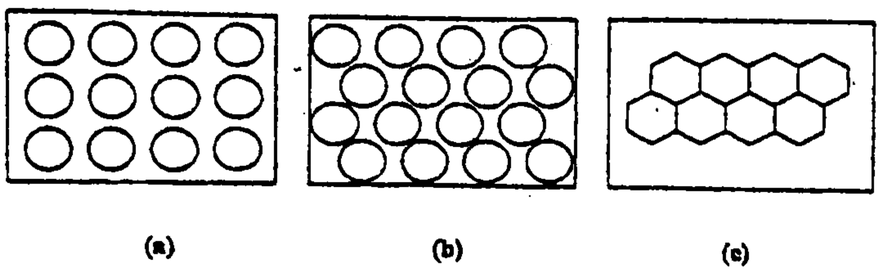


Fig. 3. Microlens packaging: (a) circular-based microlenses in rectangular net packaging, (b) honeycomb-packed circular-based microlenses, (c) honeycomb-packed hexagonal-based microlenses.

$$x_k^2 = \frac{a_k(z_P + S) + x_P f_1}{f_1 + S + z_P},$$

$$y_k^2 = \frac{b_k(z_P + S) + y_P f_1}{f_1 + S + z_P},$$

$$z_k^2 = \frac{-z_P f_1 + z_P S + S^2}{f_1 + S + z_P}. \tag{2}$$

Each intermediary voxel P_k^2 is collected by the corresponding output macrolens k and imaged at the location $P'(x_P, y_P, -z_P)$ in space, which is the equal conjugate image location of point P with respect to the Oxy plane. It does not therefore depend on the macrolens k, and the result is an integration of directional information about P that is the product of all the macrolenses and microlenses that have imaged it.

The intensity distributions related to the integral image P' of P may be recorded on a photographic plate that lies behind the recording microlens array (containing microlenses of focal distance F, situated at a distance D from the central Oxy plane; Fig. 1). P' has recorded intensity distributions in a number of microlenses, the number being determined by the axial position of P'.

If microlens k, of center $C_k(\alpha_k, \beta_k, D)$, "sees" the point P', then the recorded data behind microlens k is centered at the point P_k', whose coordinates are

$$x_k' = \alpha_k - \frac{F(x_P - \alpha_k)}{D - |z_P|},$$

$$y_k' = \beta_k - \frac{F(x_P - \beta_k)}{D - |z_P|},$$

$$z_k' = D + F. \tag{3}$$

The disparity between two recorded intensity distributions of the point P corresponding to microlenses k and j, of centers $C_k(\alpha_k, \beta_k, D)$ and $C_j(\alpha_j, \beta_j, D)$, is

$$d_{yk} = \left| \frac{D + F - |z_P|}{D - |z_P|} \right| \delta_{jk}, \tag{4}$$

where $\delta_{jk} = [(\alpha_j - \alpha_k)^2 + (\beta_j - \beta_k)^2]^{1/2}$ is the distance between C_j and C_k.

Equation (4) allows the recovery of the physical point's position when two of its intensity distributions, centered at $P_j'(x_j', y_j', D + F)$ and $P_k'(x_k', y_k', D + F)$, are known (and therefore the disparity d_{jk} between them is known). That is,

$$x_P = \frac{\alpha_k d_{jh} - x_k' \delta_{jh}}{d_{jh} - \delta_{jk}},$$

$$y_P = \frac{\beta_k d_{jk} - x_k' \delta_{jk}}{d_{jk} - \delta_{jk}},$$

$$z_P = \frac{F \delta_{jk}}{d_{jk} - \delta_{jk}} - D. \tag{5}$$

Another important consideration, particularly with respect to the way information is distributed, is the number of recorded intensity distributions of the optical reconstruction P' of a physical point P [Fig. 2(b)]. Computations show that this number depends on both the lateral position and the depth of the point, as well as on the configuration of the camera (microlens pitch, microlens type, and packaging). The packaging types are shown in Fig. 3. The maximum recorded image number for a given depth is attained for points having a central position in the imaged scene. For circular-based microlens arrays packed as a rectangular net [Fig. 3(a)], this maximum is

$$N_{\max} = \left\{ \left[ap \frac{D - |z_P|}{\phi_r(|z_P - S)} \right] + 1 \right\}^2, \tag{6}$$

and for circular- or hexagonal-based honeycomb-packed microlens configurations [Figs. 3(b) and 3(c)], it is given by

$$N_{\max} = \left\{ \left[ap \frac{D - |z_P|}{\frac{\sqrt{3}}{2} \phi_r(|z_P| - S)} \right] + 1 \right\}^2, \tag{7}$$

where ap is the side of the quasi-square camera aperture.

Comparing the results given in Eqs. (6) and (7), we can infer that given the same camera aperture, microlens pitch, and point position, the recording process using honeycomb-packed microlenses retains more angular information than that corresponding to rectangularly packed microlenses. This is due to the fact that the honeycomb packaging reduces the effective microlens pitch ϕ_r to $\sqrt{3}\phi_r/2$.

It can be seen that the number of recorded intensity-distribution components depends essentially on the camera aperture ap with marginal points having fewer encoded spatial-distribution components owing to the finite dimensions of the recording screen. In Fig. 4, a Matlab application has been used to plot the step functions that describe the dependence between the number of recorded intensity-distribution components of a point and the depth for different lateral positions of the point. Figures 4(a) and 4(b) depict the number of recorded components for points in a plane parallel to Oyz and to Oxz, respectively, with image-plane depth ranging between 0.4 and 0.58 m when the recording screen is at a distance D = 0.58 m from the central plane of a camera having an aperture ap = 0.3 m and honeycomb-packed circular or hexagonal microlenses of pitch ϕ_r = 1.65 mm in the recording array. The number of recorded intensity distributions for points having the same depth $|z| = z_0$ = 0.50 m and different lateral positions is represented in Fig. 4(c). Similarly shaped graphs are obtained in the case of circular-based microlenses packed in a rectangular net. As can be seen, points in the proximity of the recording plane are not recorded. This singularity is explained in terms of physical optics in Section 4.

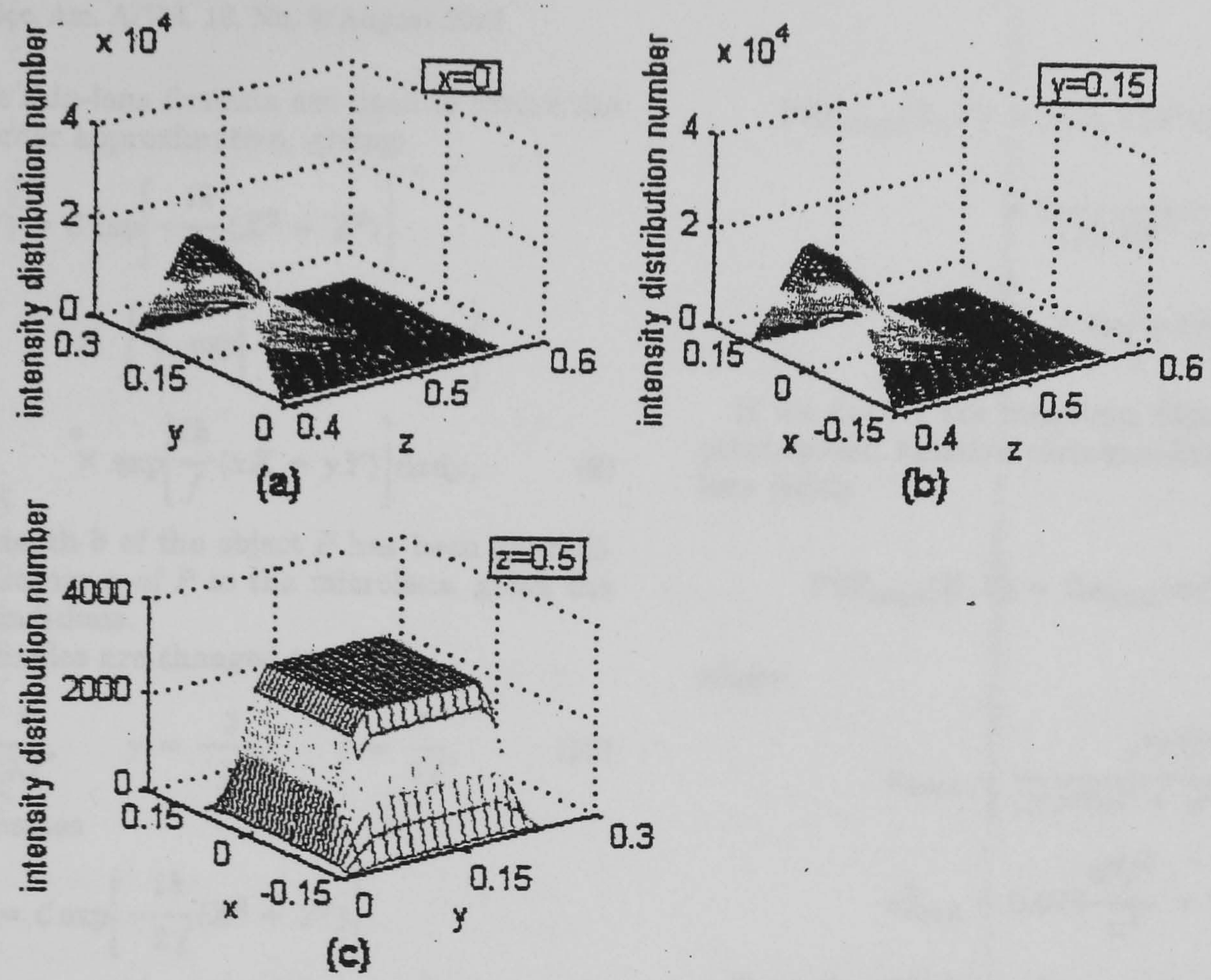


Fig. 4. Number of recorded microintensity distributions for a hemispherical microlens system: (a) x = const., (b) y = const., (c) z = const.

4. POINT-SPREAD FUNCTION IN THE THREE-DIMENSIONAL INTEGRAL RECORDING SYSTEM

The first mathematical treatment of a single-tier integral imaging system was presented by Burckhardt in the context of both remote[3] and close[4] imaging. The calculations were performed for a square-based microlens in a direction parallel to a side of the square lenslet. The present approach considers circular- and hexagonal-based spherical microlens arrays, all calculations being performed two dimensionally on the microlens aperture. To derive the point-spread function, we consider a quasi-pointwise source of light P situated at a distance a in front of a microlens (see Fig. 5). The half "thickness" of this light source is denoted by b.

The image formation is affected by so-called close imaging when the microlens screen is located in the immediate vicinity of the object point (such that a and b are comparable), whereas remote imaging operates when the distance of the imaged object to the microlens screen is relatively large compared with the object thickness ($a \gg b$). The integral image formation and recording of 3D scenes combines both remote and close imaging; therefore the analysis of each is necessary.

A. Point-Spread Function in an Integral System Based on Circular-Based Microlenses

Initially, we derive the point-spread function for a camera configuration based on circular-based microlenses, assuming that close imaging conditions exist for both image formation and recording. The plane where the spread

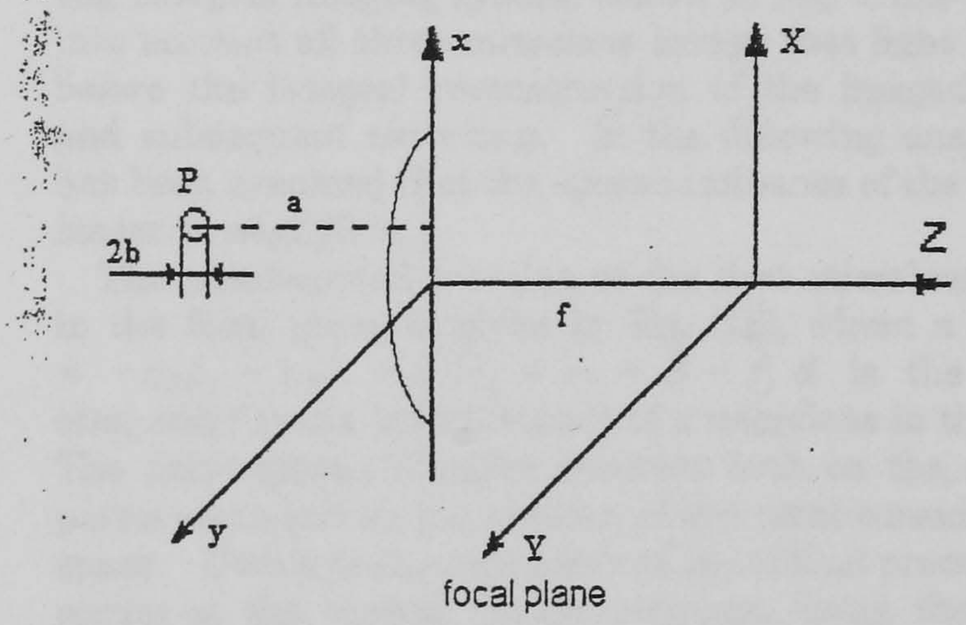


Fig. 5. Close and remote imaging.

function is to be considered is the focal plane OXY of the microlens (Fig. 5), where c denotes the image distance of the point source and f the focal distance of the microlens. The amplitude of the electric field in the focal plane is given by the Fresnel–Kirchhoff integral equation[4,11]:

$$A(X, Y) = C \int\int_{\mathcal{D}} \exp\left[\frac{ik}{2c}(x^2 + y^2)\right] \exp(-ikR)\,dx\,dy, \tag{8}$$

where $\mathcal{D} = \{(x, y) | x^2 + y^2 \leqslant r^2\}$ is the lens aperture, $R = f^2 + (X - x)^2 + (Y - y)^2$ is the distance between a point in the plane of the lens and a point in the focal plane, and $k = 2\pi/\lambda$ is the wave number. The Taylor ex-

pansion and the thin-lens formula are used to derive the following first-order approximation, giving

$$A(X, Y) = C\exp\left[-\frac{ik}{2f}(X^2 + Y^2)\right] \times \int\int_{\mathcal{D}} \exp\left[-\frac{ik}{2a}(x^2 + y^2)\right] \times \exp\left[\frac{ik}{f}(xX + yY)\right]dx\,dy, \quad (9)$$

where the half-depth b of the object P has been approximated to the distance a of P to the microlens, given the close-imaging conditions.

When the variables are changed to

$$\hat{x} = \frac{x}{\lambda f}, \quad y = \frac{y}{\lambda f}, \quad \ldots \quad \hat{r} = \frac{r}{\lambda f}, \quad (10)$$

the equation becomes

$$A(X, Y) = C\exp\left[-\frac{ik}{2f}(X^2 + Y^2)\right] \times \int\int_{\hat{\mathcal{D}}} \exp\left[-\frac{ik\lambda^2 f^2}{2a}(\hat{x}^2 + \hat{y}^2)\right] \times \exp[-i2\pi(\hat{x}X + \hat{y}Y)]d\hat{x}\,d\hat{y}, \quad (11)$$

where $\hat{\mathcal{D}}$ is the disk $\{(\hat{x}, \hat{y})|\hat{x}^2 + \hat{y}^2 \leq \hat{r}^2\}$.

It can be noticed that the latter integral represents the Fourier transform of the product between $\psi_1 = \exp[-(ik\lambda^2 f^2/2a)(\hat{x}^2 + \hat{y}^2)]$ and the characteristic function ψ_2 of the disk $\hat{\mathcal{D}}$; therefore the amplitude A is the convolution of the Fourier transforms of ψ_1 and ψ_2:

$$A(X, Y) = C\exp\left[-\frac{ik}{2f}(X^2 + Y^2)\right]\mathcal{F}(\psi_1) * \mathcal{F}(\psi_2), \quad (12)$$

where the asterisk represents convolution.

The two Fourier transforms are

$$\mathcal{F}(\psi_1) = \frac{a}{\lambda f^2}\exp\left[-\frac{i\pi a}{\lambda f^2}(X^2 + Y^2)\right], \quad (13)$$

$$\mathcal{F}(\psi_2) = C\frac{J_1(2\pi\rho\hat{r})}{2\pi\rho\hat{r}}, \quad (14)$$

where $\rho^2 = X^2 + Y^2$ and J_1 is the first-order Bessel function. To perform the convolution, we consider the approximation $2J_1(w)/w \simeq \exp(-w^2/6.5)$.

Straightforward calculations yield

$$A(\rho) = C\frac{a}{(6.0736\hat{r}^2 + i\pi a/\lambda f^2)^{1/2}}\exp\left(-\frac{ik}{2f}\rho^2\right) \times \exp\left[-\beta\rho^2\left(\frac{\pi a}{\lambda f^2} + i6.0736\hat{r}^2\right)\right], \quad (15)$$

where

$$\beta = \frac{6.0736\hat{r}^2\pi a\lambda f^2}{36.89\lambda^2 f^4\hat{r}^4 + \pi^2 a^2}. \quad (16)$$

Hence the point-spread function is given by

$$\mathrm{PSF}_{\mathrm{circ,cl}}(X, Y) = A(X, Y)A^*(X, Y) = C\frac{a^2}{(36.89\hat{r}^4 + \pi^2 a^2/\lambda^2 f^4)^{1/2}} \times \exp\left[-2\beta\frac{\pi a}{\lambda f^2}(X^2 + Y^2)\right]. \quad (17)$$

If we denote the microlens diameter by $d = 2r$, the point-spread function corresponding to a circular microlens yields

$$\mathrm{PSF}_{\mathrm{circ,cl}}(X, Y) = C\alpha_{\mathrm{circ,cl}}\exp\left(-\frac{X^2 + Y^2}{u^2_{\mathrm{circ,cl}}}\right), \quad (18)$$

where

$$\alpha_{\mathrm{circ,cl}} = \frac{a^2\lambda^2 f^2}{(2.305d^4 + \pi^2 a^2\lambda^2)^{1/2}} \quad (19)$$

$$u^2_{\mathrm{circ,cl}} = 0.076\frac{d^2 f^2}{a^2} + 0.328\frac{\lambda^2 f^2}{d^2}. \quad (20)$$

From Eq. (20) it can be seen that the first term of the sum expresses the defocus influence and the second expresses the diffraction. In addition, both $\alpha_{\mathrm{circ,cl}}$ and $u_{\mathrm{circ,cl}}$ contain depth information about the imaged point.

The calculation of the point-spread function of the two-tier integral imaging system shown in Fig. 1 has to take into account all three microlens arrays that light crosses before the integral reconstruction of the imaged object and subsequent recording. In the following analysis it has been assumed that the spread influence of the macrolenses is negligible.

The point-spread function of the first microlens array in the focal plane is given by Eq. (18), where $a = a_{\mathrm{ATS}} = -z_P f_1 + z_P S + S^2/f_1 + z_P + S - f$, d is the diameter, and f is the focal distance of a microlens in the ATS. The point-spread function depends both on the camera parameters and on the position of the point considered in space. Owing to the symmetry of the optical process that occurs at the central double-microlens array, the point-spread function of the second microlens array is also given by Eq. (18). Therefore the spread function of the spatial inversion process (associated with the central screen) is given by the convolution of Eq. (18) with itself:

$$\mathrm{PSF}^{\mathrm{ATS}}_{\mathrm{circ,cl}} = \mathrm{PSF}^{\mathrm{ATS}}_{\mathrm{circ,cl}} * \mathrm{PSF}^{\mathrm{ATS}}_{\mathrm{circ,cl}}. \quad (21)$$

A straightforward computation yields

$$\mathrm{PSF}^{\mathrm{ATS}}_{\mathrm{circ,cl}}(X, Y) = C(\alpha^{\mathrm{ATS}}_{\mathrm{circ,cl}})^2 u^{\mathrm{ATS}}_{\mathrm{circ,cl}}\exp\left[-\frac{X^2 + Y^2}{2(u^{\mathrm{ATS}}_{\mathrm{circ,cl}})^2}\right]. \quad (22)$$

The point-spread function associated with the recording microlens screen in the focal plane (which is the plane of the photographic plate) is analogous to Eq. (18) but contains different parameters and is given by

$$\mathrm{PSF}^{\mathrm{rec}}_{\mathrm{circ,cl}}(X, Y) = C\alpha^{\mathrm{rec}}_{\mathrm{circ,cl}}\exp\left[-\frac{X^2 + Y^2}{(u^{\mathrm{rec}}_{\mathrm{circ,cl}})^2}\right], \quad (23)$$

where $\alpha^{rec}_{circ,cl}$ and $u^{rec}_{circ,cl}$ have expressions similar to Eqs. (19) and (20), where $d := d_r$, $f := F$ are the diameter and the focal distance respectively, of the recording microlens screen, and $a := a_r$ is the distance between the reconstructed model P' and the recording microlens array ($a_r = D - |z_P|$).

The spread function of the entire optical process combines the spread effect of both the central double-microlens screen and the recording array. Therefore it is the convolution of $PSF^{ATS}_{circ,cl}$ and $PSF^{rec}_{circ,cl}$:

$$PSF^{total}_{circ,cl} = PSF^{ATS}_{circ,cl} * PSF^{rec}_{circ,cl}. \tag{24}$$

Computations yield

$$PSF^{total}_{circ,cl}(X, Y) = C(\alpha^{ATS}_{circ,cl})^2 \alpha^{rec}_{circ,cl} \frac{(u^{ATS}_{circ,cl})^2 u^{rec}_{circ,cl}}{[2(u^{ATS}_{circ,cl})^2 + (u^{rec}_{circ,cl})^2]^{1/2}} \times \exp\left[-\frac{X^2 + Y^2}{2(u^{ATS}_{circ,cl})^2 + (u^{rec}_{circ,cl})^2}\right]. \tag{25}$$

Equation (25) expresses the spread function in local coordinates associated with the microlens recording screen. Using the global Cartesian system depicted in Fig. 1, we can write the point spread behind a specific microlens k of the recording array that images a point P as

$$PSF^{total,k}_{circ,cl}(x, y) = C(\alpha^{ATS}_{circ,cl})^2 \alpha^{ATS}_{psh,cl} \frac{(u^{ATS}_{circ,cl})^2 u^{rec}_{circ,cl}}{[2(u^{ATS}_{circ,cl})^2 + (u^{rec}_{circ,cl})^2]^{1/2}} \times \exp\left[-\frac{(x - x_k')^2 + (y - y_k')^2}{2(u^{ATS}_{circ,cl})^2 + (u^{rec}_{circ,cl})^2}\right], \tag{26}$$

where x_k', and y_k' are the coordinates of the recorded intensity distribution P_k' behind microlens k, given in Section 3.

Parallax information about any imaged point P is distributed over the photographic plate as a set of intensity distributions recorded in a number of microlenses, the number being determined by the axial position of the optical reconstruction P'. The point-spread function of the whole process is the sum of these disparate distributions:

$$PSF^{global}_{circ,cl}(x, y) = \sum_k PSF^{total,k}_{circ,cl}(x, y), \tag{27}$$

where the summation is carried out with respect to the indices k of the microlenses involved in recording point P.

When remote imaging is assumed, a similar procedure for deriving the point-spread function is used.

The laws of geometrical optics show that the image distance is obtained as

$$c = f + \frac{f^2}{a^2} b,$$

where a is the distance between object and lens. With this equation, the Fresnel–Kirchhoff integral (8) becomes

$$A_{circ,rm}(X, Y) = C \int\int_D \exp\left[-\frac{ik}{2a^2}(x^2 + y^2)\right] \times \exp\left(\frac{ik}{f}(xX + yY)\right) dx\,dy. \tag{28}$$

Computations similar to those performed in the close-imaging case lead to the following expression for the point-spread function for a circular microlens:

$$PSF_{circ,rm}(X, Y) = C\alpha_{circ,rm} \exp\left(-\frac{X^2 + Y^2}{u^2_{circ,rm}}\right), \tag{29}$$

where

$$\alpha_{circ,rm} = \frac{\lambda^2 a^4 f^2}{b(2.305 d^4 b^2 + \pi^2 \lambda^6 f^4 a^4)^{1/2}}, \tag{30}$$

$$u^2_{circ,rm} = 0.076 \frac{d^2 f^2 b^2}{a^4} + 0.328 \frac{\lambda^2 f^2}{d^2}. \tag{31}$$

The point spread due to the whole optical process in the focal plane of a recording microlens k expressed in local coordinates attached to the microlens is

$$PSF^{total,k}_{circ,rm}(X, Y) = C(\alpha^{ATS}_{circ,rm})^2 \alpha^{rec}_{circ,rm} \frac{(u^{ATS}_{circ,rm})^2 u^{rec}_{circ,rm}}{[2(u^{ATS}_{circ,rm})^2 + (u^{rec}_{circ,rm})^2]^{1/2}} \times \exp\left[-\frac{X^2 + Y^2}{2(u^{ATS}_{circ,rm})^2 + (u^{rec}_{circ,rm})^2}\right], \tag{32}$$

where $\alpha^{ATS}_{circ,rm}$, $\alpha^{rec}_{circ,rm}$, and $u^{ATS}_{circ,rm}$, $u^{rec}_{circ,rm}$ are given by expressions analogous to Eqs. (30) and (31), and the global point-spread function for remote imaging conditions,

$$PSF^{global}_{circ,rm}(x, y) = \sum_k PSF^{total,k}_{circ,rm}(x, y), \tag{33}$$

is analogous to Eq. (27).

B. Point-Spread Function in an Integral System Based on Hexagonal Microlenses

We can make analogous analyses in order to derive the point-spread function for camera configurations that involve hexagonal microlenses in both close- and remote-imaging contexts.

If we consider a hexagonal microlens of side l and perform similar computations, the expression for the amplitude is given by Eq. (12). The salient difference is that ψ_2 in this case represents the Fourier transform of the characteristic function of a hexagonal domain of side $\hat{l} = l/\lambda f$, where f is the focal distance. This Fourier transform is real owing to the symmetry of ψ_2 and is given by

$$\mathcal{F}(\psi_2) = \frac{1}{2\pi^2 Y(X + Y/\sqrt{3})} \times \{\cos[\pi\hat{l}(Y - X\sqrt{3})] - \cos(2\pi\hat{l}Y)\} + \frac{1}{2\pi^2 Y(X - Y/\sqrt{3})}\{\cos(2\pi\hat{l}Y) - \cos[\pi\hat{l}(Y + X\sqrt{3})]\}. \quad (34)$$

For the convolution to be performed, an approximation to this particular Fourier transform is necessary. The graphical representation of Eq. (34), which has the shape of an asymmetric bell, suggests a Gaussian approximation, which results in

$$\mathcal{F}(\psi_2) \simeq \frac{3\sqrt{3}}{2}\hat{l}^2 \exp[(-4.8\hat{l}^2X^2 - 5.3\hat{l}^2Y^2)]. \quad (35)$$

Effecting the approximation, the following expression for the point-spread function of a hexagonal (micro)lens under close imaging conditions,

$$\mathrm{PSF}_{\mathrm{hex,cl}}(X, Y) = C\alpha_{\mathrm{hex,cl}} \exp\left(-\frac{X^2}{u^2_{\mathrm{hex,cl}}} - \frac{Y^2}{v^2_{\mathrm{hex,cl}}}\right), \quad (36)$$

is obtained, where

$$\alpha_{\mathrm{hex,cl}} = \frac{a^2l^2}{f^2[(23.04l^4 + \pi^2a^2\lambda^2)(28.09l^4 + \pi^2a^2\lambda^2)]^{1/2}}, \quad (37)$$

$$u_{\mathrm{hex.cl}} = 0.243\frac{l^2f^2}{a^2} + 0.104\frac{\lambda^2f^2}{l^2}, \quad (38)$$

$$v_{\mathrm{hex.cl}} = 0.270\frac{l^2f^2}{a^2} + 0.094\frac{\lambda^2f^2}{l^2}. \quad (39)$$

With this expression, the point-spread function of the entire optical process of the camera with hexagonal microlenses yields

$$\mathrm{PSF}^{\mathrm{total},k}_{\mathrm{hex,cl}}(X, Y) = C(\alpha^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2\alpha^{\mathrm{rec}}_{\mathrm{hex,cl}}\Gamma_{\mathrm{hex,cl}} \times \exp\left(-\frac{X^2}{U^2_{\mathrm{hex,cl}}} - \frac{Y^2}{V^2_{\mathrm{hex,cl}}}\right), \quad (40)$$

where

$$\Gamma_{\mathrm{hex,cl}} = \frac{(u^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2(v^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2u^{\mathrm{rec}}_{\mathrm{hex,cl}}v^{\mathrm{rec}}_{\mathrm{hex,cl}}}{\{[2(u^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2 + (u^{\mathrm{rec}}_{\mathrm{hex,cl}})^2][2(v^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2 + (v^{\mathrm{rec}}_{\mathrm{hex,cl}})^2]\}^{1/2}}, \quad (41)$$

$$U^2_{\mathrm{hex,cl}} = 2(u^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2 + (u^{\mathrm{rec}}_{\mathrm{hex,cl}})^2, \quad (42)$$

$$V^2_{\mathrm{hex,cl}} = 2(v^{\mathrm{ATS}}_{\mathrm{hex,cl}})^2 + (v^{\mathrm{rec}}_{\mathrm{hex,cl}})^2, \quad (43)$$

and the α, u, and v factors are given by Eqs. (37) and (38), whose parameters have been adapted for the corresponding microlens screen, as indicated by the indices rec and ATS.

Analogously, the total point-spread function when remote imaging is assumed is given by

$$\mathrm{PSF}^{\mathrm{total},k}_{\mathrm{hex,rm}}(X, Y) = C(\alpha^{\mathrm{ATS}}_{\mathrm{hex,rm}})^2\alpha^{\mathrm{rec}}_{\mathrm{hex,rm}}\Gamma_{\mathrm{hex,rm}} \times \exp\left(-\frac{X^2}{U^2_{\mathrm{hex,rm}}} - \frac{Y^2}{V^2_{\mathrm{hex,rm}}}\right), \quad (44)$$

where $\Gamma_{\mathrm{hex,rm}}$, $U_{\mathrm{hex,rm}}$, and $V_{\mathrm{hex,rm}}$ correspond to Eq. (41), and

$$\alpha_{\mathrm{hex,rm}} = \frac{a^4l^4\lambda^2}{\sqrt{23.04l^4 + \pi^2\lambda^6a^4f^4}\sqrt{28.09l^4 + \pi^2\lambda^6a^4f^4}}, \quad (45)$$

$$u_{\mathrm{hex,rm}} = 0.243\frac{l^2f^2b^2}{a^4} + 0.104\frac{\lambda^2f^2}{l^2}, \quad (46)$$

$$v_{\mathrm{hex,rm}} = 0.270\frac{l^2f^2b^2}{a^4} + 0.094\frac{\lambda^2f^2}{l^2}. \quad (47)$$

The global point-spread function for a hexagonal microlens camera configuration is given, for both close and remote imaging, by the sum of local point-spread functions behind all microlenses that image the quasi-pointwise light source. The expression is simply

$$\mathrm{PSF}^{\mathrm{global}}_{\mathrm{hex}}(x, y) = \sum_k \mathrm{PSF}^{\mathrm{total},k}_{\mathrm{hex}}(x, y). \quad (48)$$

C. Spread versus Depth in Integral Imaging

The shape of $\mathrm{PSF}^{\mathrm{total},k}$, which is a two-variable Gaussian function, symmetric for circular and asymmetric for hexagonal microlenses, suggests that the energy forming the recorded intensity distribution is concentrated on a circular or elliptic spot whose dimensions depend on point depth. Matlab applications were used to plot the dependence on depth for the dimensions (diameter or axes) of a circular or elliptic spot that concentrates more than 98% of the recorded intensity. Both close- and remote-imaging conditions were studied. Tables 1 and 2 present spread dependence on depth in a camera configuration

Table 1. Spread versus Depth at Close Imaging for Circular- and Hexagonal-Based Microlenses[a]

Point Depth $\lvert z\rvert$ (m)	Spread at Close Imaging (mm)		$\mathrm{Spread}_{\mathrm{hex}}/\mathrm{Spread}_{\mathrm{circ}}$[b]
	Circular Lens	Hexagonal Lens	
0.31	0.0120	0.0126	1.043
0.34	0.0133	0.0140	1.045
0.37	0.0151	0.0158	1.047
0.40	0.0173	0.0182	1.048
0.43	0.0205	0.0216	1.050
0.46	0.0253	0.0266	1.051
0.49	0.0331	0.0348	1.052
0.52	0.0481	0.0507	1.053
0.55	0.0889	0.0937	1.053
0.58	0.0621	0.0655	1.053

[a] The depth of the recording screen is $D = 0.585$ m.
[b] Ratio between the spread for a hexagonal and for a circular microlens system.

Table 2. Spread versus Depth at Remote Imaging for Circular- and Hexagonal-Based Microlenses[a]

Point Depth $\|z\|$ (m)	Spread at Remote Imaging (mm)		$\text{Spread}_{hex}/\text{Spread}_{circ}$[b]
	Circular Lens	Hexagonal Lens	
0.36	0.004006	0.00363847	0.908
0.39	0.004007	0.00363848	0.907
0.43	0.004009	0.00363851	0.907
0.46	0.004014	0.00363855	0.906
0.47	0.004016	0.00363857	0.905
0.48	0.004020	0.00363859	0.905
0.49	0.004026	0.00363862	0.903
0.50	0.004035	0.00363867	0.901
0.51	0.004051	0.00363873	0.898
0.52	0.004081	0.00363883	0.891

[a] The depth of the recording screen is $D = 0.585$ m.
[b] Ratio between the spread for a hexagonal and for a circular microlens system.

corresponding to a recording screen position $D = 0.585$ m, at close imaging (Table 1) and remote imaging (Table 2). A wavelength $\lambda = 500$ nm has been assumed. The tables show that the spread is larger when the recorded points are nearer the recording screen. For points reconstructed in the immediate vicinity of the recording plane at a distance smaller than the focal distance of a recording microlens, the spread tends to infinity; therefore these points do not contribute *stricto sensu* to the recorded spatial information. This result confirms that the number of contributing intensity distributions for these points is zero, as proved on a geometrical basis in Section 3. A continuous decrease of the spread is noticed when the distance from the optical reconstruction to the recording screen grows. Hence the spread dimensions encode depth information about the imaged scene.

The spread-versus-depth relationship for hexagonal and circular microlens systems is presented for close and remote imaging, respectively, in Tables 1 and 2. The ratio between the spread for the two types of microlenses as a function of depth is presented in the last columns of these tables. The same microlens pitches have been considered for both circular and hexagonal microlenses ($\phi_{ATS} = 0.125$ mm, $\phi_{rec} = 1.206$ mm). Making a comparative analysis between circular and hexagonal microlenses from this point of view, we can observe that in the case of close imaging, the spread for a hexagonal microlens array is 4–5% larger than that of the corresponding circular microlens array. However, in the case of remote imaging, the spread for a hexagonal microlens system is less than 90% of that of a system based on circular microlenses, and for points situated at a distance of 3 cm in front of the recording screen (which can still be considered to fulfill remote imaging conditions) it is reduced twofold.

Since the recording of deep scenes involves remote imaging for depth planes situated further from the recording screen and close imaging for the proximate depth range, the hexagonal honeycomb-packed microlens system can be considered to have a better performance than the similarly packed circular-based microlens system. Additionally, the hexagonal microlenses ensure an efficient use of the recording surface, such a structure eliminating the inactive areas between microlenses that appear when the operation is with a circular-based microlens configuration.

5. CONCLUSIONS

A model of the 3D integral recording camera system, involving both geometrical and physical aspects of the optical process, has been described. Two variants of the camera, one based on circular and another based on hexagonal microlenses, have been analyzed and compared in terms of image formation and point spread. The analysis has taken into account the combined effect of all the microlenses of the arrays in the image formation and has shown that the recording process that uses honeycomb-packed hexagonal microlenses retains more angular details than that corresponding to rectangularly packed circular-based microlens array. Moreover, the hexagonal microlens arrays do not have inactive areas and ensure an efficient use of the recording surface. The spreads of the two systems for close imaging are similar, but the hexagonal microlens configuration has a lower point spread for remote imaging situations, which is obviously a very important feature in the recording of deep scenes.

Address correspondence to Silvia Manolache at the address on the title page or by e-mail, silvia@dmu.ac.uk.

REFERENCES

1. G. Lippmann, "La photographie intégrale," C. R. Hebd. Seances Acad. Sci. **146**, 446–451 (1908).
2. H. E. Ives, "Optical properties of a Lippmann lenticulated sheet," J. Opt. Soc. Am. **21**, 171–176 (1931).
3. C. Burckhardt, "Optimum parameters and resolution limitation of integral photography," J. Opt. Soc. Am. **58**, 71–76 (1968).
4. C. Burckhardt, R. J. Collier, and E. T. Doherty, "Formation and inversion of pseudoscopic images," Appl. Opt. **7**, 627–631 (1968).
5. T. Okoshi, *Three Dimensional Imaging Techniques* (Academic, London, 1976).
6. N. Davies, M. McCormick, and L. Yang, "Three-dimensional imaging systems: a new development," Appl. Opt. **27**, 4520–4528 (1988).
7. F. Okano, J. Arai, H. Hoshino, and I. Yuyama, "Three-dimensional video system based on integral photography," Opt. Eng. **38**, 1072–1077 (1999).
8. M. Brewin, M. Forman, and N. Davies, "Electronic capture and display of full parallax 3D images," in *Stereoscopic Displays and Virtual Reality Systems II*, S. S. Fisher, J. O. Merritt, and M. T. Bolas, eds., Proc. SPIE **2409**, 118–124 (1995).
9. N. Davies and M. McCormick, "Holoscopic imaging with true 3D-content in full natural colour," J. Phot. Sci. **40**, 46–49 (1992).
10. N. Davies, M. McCormick, and M. Brewin, "Design and analysis of an image transfer system using microlens arrays," Opt. Eng. **33**, 3624–3633 (1994).
11. M. V. Klein and T. E. Furtak, *Optics* (Wiley, New York, 1986).

Journal of VLSI

SIGNAL PROCESSING SYSTEMS

for Signal, Image, and Video Technology

KLUWER ACADEMIC PUBLISHERS
Boston/U.S.A. Dordrecht/Holland London/U.K.

3D-object space reconstruction from planar recorded data of 3D-integral images

Silvia Manolache[1,2*], S.Y. Kung[1], Malcolm McCormick[2],
Amar Aggoun[2]
[1] *Princeton University, Department of Electrical Engineering, Princeton NJ 08540, USA, E-mail: kung@ee.princeton.edu*
[2] *De Montfort University, Department of Engineering and Technology, The Gateway, Leicester LE1 9BH, UK, Tel: +44-(0)116-2551551x8082, Fax: +44-(0)116-2577692, E-mail: silvia@dmu.ac.uk*

Abstract. The paper presents a novel algorithm for object space reconstruction from the planar (2D) recorded data set of a 3D-integral image. The integral imaging system is described and the associated point spread function is given. The space data extraction is formulated as an inverse problem, which proves ill-conditioned, and tackled by imposing additional conditions to the sought solution. An adaptive constrained 3D-reconstruction regularization algorithm based on the use of a sigmoid function is presented. A hierarchical multiresolution strategy which employes the adaptive constrained algorithm to obtain highly accurate intensity maps of the object space is described. The depth map of the object space is extracted from the intensity map using a weighted Durbin–Willshaw algorithm. Finally, illustrative simulation results are given.

Keywords: integral imaging, object space reconstruction, inverse problems, regularization methods, gradient descent, Durbin–Willshaw scheme

1. Introduction

The development of 3D-imaging systems has been a constant pursuit of the scientific as well as of the entertainment community in the new technological era. Integral photography was pioneered by Lippmann ((Lippmann, 1908), 1908), who used microlens arrays to create, record on photographic film, and replay integral three dimensional images. Since 1908, when it was first reported, the integral photographic technique has been improved as a result of theoretical studies, technical innovations of the optical systems, and progress in microlens manufacturing. Integral imaging resembles to holography, but it uses natural light and reproduces true colour optical models. Hence, it offers a viable alternative to other autostereoscopic systems ((Okoshi, 1976)).

* The first author performed the work while at Princeton University as a Visiting Research Student, under the frame of the Exchange Visitor Programme P1 180.

The optics of an advanced form of integral imaging system – employing a two tier optical network – in which a true 3D optical reconstruction of a scene is transferred to the capture device, has been described in detail by Davies and McCormick ((Davis and McCormick, 1992), (Davis and McCormick, 1994)).

Extracting depth information from 3D-integral images has various applications, which range from remote inspection in robotic vision and medical imaging to combining real and computer generated 3D-pictures in a virtual studio. The aim of the present paper is to describe a method of reconstructing the composition of a 3D-object space from 2D-recorded data encoding the scene. The algorithm is applicable both to integral imaging and holography. The current work has dwelt upon the application to integral imaging. More specifically, the depth estimation from 3D-integral pictures is formulated as an inverse problem: given the image, i.e. the recorded 'effect' of the object space, find the 'cause' which had produced it, i.e. recover the intensity map and, consequently, the composition and the depth of the object scene. The direct problem - image formation and recording - has been studied in a previous paper, and the equation of the point spread function for the lenticular integral imaging system has been derived ((Manolache, Aggoun et alt., 1999)). Inverse problems in imaging are ill-posed and their discrete correspondents are ill-conditioned due to the inherent loss of information associated with the direct process ((Bertero and Boccacci, 1998)). In order to cure the ill-posedness of the problem, approximate solutions satisfying additional constraints coming from the physics of the problem are searched.

The present work comprises an adaptive regularization scheme for obtaining a constrained least squares solution of the depth extraction problem, which is hierarchically applied in order to obtain high resolution object space reconstruction. The output of the hierarchical algorithm is an *intensity map* of the object space. Using a weighted Durbin–Willshaw scheme, the *depth map* (contour) of the object space is extracted from the intensity map.

Organization of the paper. The integral imaging system and the associated point spread function are presented in section 2. The formulation of the object reconstruction as an inverse problem is given in section 3. The adaptive constrained 3D-reconstruction (AC3DR) algorithm is presented in section 4 and the hierarchical scheme which employes AC3DR to produce the intensity map of the object space is described in section 5. The depth map generation using a weighted Durbin-Willshaw algorithm is given in section 6. Section 7 contains illustrative simulation results.

2. 3D-integral imaging system and associated point spread function

The optical arrangement of the 3D-integral imaging system, shown in Figure 1, comprises two macrolens arrays (MA1 and MA2) placed equidistantly behind and in front of an autocollimating transmission screen (ATS). The ATS is made up of two microlens arrays separated by their joint focal distance. The recording plane, a photographic plate whose position coincides with the focal plane of a microlens array (RA), lies within the optical model.

The paper considers a camera configuration where the ATS contains square based hemispherical lenslets and the recording array is made up of identical cylindrical surface microlenses. This system is known as a *3D-lenticular integral imaging system* ((Davis and McCormick, 1992), (Manolache, Aggoun et alt., 1999)) and produces 3D-images containing horizontal parallax.

The object is imaged by the input macrolenses which transmit compressed transposed images that occur at or near the central double microlens screen (ATS). The screen inverts the optical sense of each intermediary image (see Figure 2), and, simultaneously, presents these spatially reversed 3D-optical models to the corresponding output macrolenses. The output macrolenses retranspose the optical models to the correct spatial location. The final integrated image, formed by superposed optical models projected by the second macrolens array, is a true 3D optical 1:1 reconstruction of the object. It is recorded as a *2D-sampled data set* possessing either horizontal parallax (if cylindrical surface microlenses are used in the recording array), or continuous parallax (in the case of the spherical or square based microlens arrangement).

The integral image recorded in the focal plane of a recording microlens array as a *planar* sampled data set contains all the 3D information related to the object space. Each microlens of the recording array samples a fractional part of the scene, many microlenses recording directional information of the scene from different viewing angles. Therefore, parallax information about any particular point is spread over the recording plane. Redisplay of the full spatial model as a real 3D image can be effected by overlaying the sampled data set by an integral decoding element.

The indices of the microlenses of the recording screen which record a certain point and the positions of the recorded samples of a point can be determined. They depend on the spatial position of the recorded point and on the camera aperture. ((Manolache, Aggoun et alt., 1999)).

The spread function of the entire optical process combines the spread effect of both the central double microlens screen and the recording

array ((Manolache, Aggoun et alt., 1999)). Its component behind a microlens l which 'sees' the object point (x, y, z) is given by:

$$K_{l(x,y,z)}(X,Y) = \alpha(z)\exp\left(-\frac{(X-X_l)^2}{a_1^2}\right)\exp\left(-\frac{(Y-Y_l)^2}{a_2^2}\right), \quad (1)$$

where (X, Y) are the coordinates in the image space, (X_l, Y_l) represents the point of maximum intensity in the microimage of (x, y, z) formed behind microlens l and:

$$\begin{aligned}
\alpha(z) &= \frac{v^4 w_1^2 w_2^2}{\sqrt{(v^2+w_1^2)(v^2+w_2^2)}} \\
a_1^2 &= v^2 + w_1^2, \qquad a_2^2 = v^2 + w_2^2 \\
v^2 &= 0.212\frac{\phi^2 f^2}{a(z)^4} + 0.49\frac{\lambda^2 f^2}{\phi^2} \\
w_1^2 &= 0.106\frac{\psi_r^2 F^2}{a_r(z)^4} + 0.245\frac{\lambda^2 F^2}{\psi_r^2} \\
w_2^2 &= 0.106\frac{\phi_r^2 F^2}{a_r(z)^4} + 0.245\frac{\lambda^2 F^2}{\phi_r^2}
\end{aligned} \quad (2)$$

depend explicitly on lenslet parameters and point depth z ((Manolache, Aggoun et alt., 1999)). In the above equations, ψ, ϕ, and f are the dimensions and the focal distance of a cylindrical surface ATS microlens, ψ_r, ϕ_r, and F are the dimensions and the focal distance of a recording microlens, and λ is the wavelength. All equations refer to the coordinate system depicted in Figure 1.

The total point spread function is the sum of the local spreads behind all the microlenses which image the point (x, y, z):

$$K(X,Y,x,y,z) = \sum_{l(x,y,z)} K_{l(x,y,z)}(X,Y). \quad (3)$$

Equation (1) indicates that the spread function is space variant and that it holds depth information in all the factors. This fact will be used in recovering three dimensional data about the object space. Due to the small values of the parameters involved in the point spread function equation, only the significant factors should be considered in a computational approach of the depth extraction problem. When using lenticular arrays (i.e. cylindrical surface microlenses), the spread on the x-dimension is very small, so the x-factors in the point spread expression can be dropped.

3. Formulation of the object space reconstruction as an inverse problem

The image formation equation for the integral imaging system is:

$$g(X,Y) = \int I(x,y,z) \sum_{l(x,y,z)} K_l(X,Y,x,y,z) dx\, dy\, dz\,, \tag{4}$$

where $I(x,y,z)$ is the intensity at point (x,y,z).

It is noticed that the integral image formation process *is not a convolution process* because the critical shift invariance property does not hold. Therefore classical deconvolution methods cannot be used to tackle the problem of scene reconstruction from an integral image.

However, the point spread function provides a linear image formation operator A such that equation (4) be written as:

$$g = A\, I\,. \tag{5}$$

THEOREM 1 (UNIQUENESS OF THE SOLUTION). *The discrete equation (5) has a unique solution in the case of noise free image formation.*

Proof. In the ideal case of noise free imaging conditions, equation (5) has a solution $I^* \geq 0$. The solution I^* is assumed sparse as it corresponds to a wide space with relatively few radiating points. Supposing that there is a $\Delta I \in R^n$ such that $A\Delta I = 0$, then it is natural to assume that ΔI has frequent interleavings +/-, due to the fact that A has only positive elements. These interleavings are very likely to correspond to positions where the elements of the vector I^* are null. Therefore, $I^* + \Delta I$ has both positive and negative components, thus it cannot be a (positive) solution of the imaging equation (5). It follows that I^* is the unique solution of (5). ■

The reconstruction of the intensity distribution in the object is the inverse process of the image formation given by equation (5). The discrete variant of the problem has been treated considering various samplings of the object space. Numerical experiments have shown that the matrix associated with the point spread function is very ill-conditioned. Alternative methods of reconstructing the object space are thus necessary and one of them is presented in the following section.

4. Adaptive constrained 3D-reconstruction algorithm

Additional constraints coming from a priori knowledge about the object space have to be imposed in order to get a *regularized solution* $\hat{I}$ which corresponds to the physical reality ((Bertero and Boccacci, 1998), (Tikhonov, 1963)). More precisely, the solution $\hat{I}$ has to be such that $A\,\hat{I}$ is as close as possible to the image g, i.e.:

$$||A\,\hat{I} - g|| = minimum\,. \tag{6}$$

Also, a condition of positivity and boundedness is necessary, as there are neither negative nor infinite intensity objects:

$$\hat{I} \in [0, T]^n\,. \tag{7}$$

$\hat{I}$ satisfying (7) can be written as the thresholded value of arbitrary data u: $\hat{I} = f(u)$, where f is a thresholding function. The projection of the interval $[0, T]$ onto R is not differentiable, thus a sigmoid approximation of it, $f_\sigma : R^n \rightarrow [0, T]^n$, has been preferred:

$$f_\sigma(u_i) = \frac{T}{1 + exp\left(-\left(\frac{u_i}{\sigma} - k\right)\right)}\,,\ i = 1 \ldots n\,, \tag{8}$$

where n is the number of components of the vector $\hat{I}$, i.e. the number of sampling points chosen in the object space. σ is an arbitrary number which controls the slope of the thresholding curve and k is a translation term (Figure 3).

To sum up, the sought vector $\hat{I}$ has the form $\hat{I} = f_\sigma(u)$ such that u minimizes the discrepancy functional $\epsilon_\sigma(.) = ||A\,f_\sigma(.) - g||$.

The gradient of ϵ_σ with respect to σ is:

$$\nabla\epsilon_\sigma = (A\ f_\sigma(u) - g)^T\ A \begin{bmatrix} f'_\sigma(u_1) & 0 & \ldots & 0 \\ 0 & f'_\sigma(u_2) & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & f'_\sigma(u_n) \end{bmatrix}\,. \tag{9}$$

The variation of u which leads to a constrained least-squares solution is on the descent direction of the gradient $\nabla_\sigma\epsilon_\sigma$:

$$\Delta u = -\eta\nabla\epsilon_\sigma\ ,\eta > 0\,. \tag{10}$$

ALGORITHM 1 (Adaptive constrained 3D-reconstruction (AC3DR)). *The current solution u_m is modified according to the gradient descent rule:*

$$u_{m+1} = u_m - \eta_m\nabla\epsilon_\sigma(u_m)\ ,$$

*where η_m is chosen using a suitable steplength algorithm ((*Ortega and Rheinboldt, 1970*)) such that $\epsilon_\sigma(u_{m+1}) \leq \epsilon_\sigma(u_m)$ and:*

$$\nabla\epsilon_\sigma(u_m) = (A\ f_\sigma(u_m) - g)^T\ A \begin{bmatrix} f'_\sigma(u_{m1}) & 0 & \dots & 0 \\ 0 & f'_\sigma(u_{m2}) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & f'_\sigma(u_{mn}) \end{bmatrix}$$

The algorithm starts with an arbitrary initial solution u_0, e.g. $u_0 = 0$, and is iterated until convergence is reached. It can be proved that there exits a steplength $\eta_i > 0$ at each iteration such that the algorithm converge to a local solution for any initial value $u_0 \in (-M, k\sigma)^n$, where M is large fixed real number.

Finally, the *regularized solution* given by the AC3DR algorithm is *thresholded automatically* using a 3-equidistant bin histogram. The values contained in the lowest bin, i.e. the 'background', are set to 0 and the rest of the values are kept. (Figure 4) Numerical simulations have shown that this 3 bin histogram thresholding rule provides reliable solutions for the current type of object reconstruction.

The solution obtained with the adaptive constrained algorithm is an approximation of the global unique solution I^* whose uniqueness has been stated in Theorem 1.

5. Hierarchical adaptive constrained 3D-reconstruction algorithm

The ill-conditioning of the depth extraction problem is determined by the existence of very small singular values of the operator A, which virtually increase the dimension of the null-space of A in numerical applications and make the solution of the equation (5) be very unstable. Such an ill-conditioning may be alleviated by reducing the size of the matrix corresponding to A, so by considering only a relatively low number of sampling points in the object space. This operation suffers from the drawback of inducing low resolution in the object reconstruction.

To allow finer resolution and yet not increase unduly the null-space of A, a *hierarchical multiresolution strategy* is proposed. It is based on performing localized regularization of the solution on certain regions of the object space at the current resolution and/or at increased resolution. This approach consists of the following:

ALGORITHM 2 (Hierarchical AC3DR (HAC3DR)).

1. Initial sampling and regularisation: *Sample a large domain* $D^{(1)}$ *using an initial grid* Σ_1 *and obtain a reconstruction* $\hat{I}_{D^{(1)}}$ *from the equation* $g = A_{D^{(1)}}^{\Sigma_1} \hat{I}_{D^{(1)}}^{\Sigma_1}$ *using AC3DR algorithm.*

2. Localized regularization: *Select the high intensity regions* D_1,.., D_m *of the object space from the current reconstruction* $\hat{I}_{D^{(1)}}^{\Sigma_1}$ *and consider the rest of the space as determined. Perform local regularization of the reconstruction solution on the union* $D^{(2)} = D_1 \cup \ldots \cup D_m$ *using one of the following alternatives:*

   a) Local regularization at the same resolution: *Consider the restriction of the sampling* Σ_1 *to the current domain* $D^{(2)}$ *and obtain a more accurate solution* $\hat{I}_{D^{(2)}}^{\Sigma_1}$ *from the equation* $g = A_{D^{(2)}}^{\Sigma_1} \hat{I}_{D^{(2)}}^{\Sigma_1}$ *using the AC3DR algorithm.*

   b) Local regularization at increased resolution: *Resample the current domain* $D^{(2)} = D_1 \cup \ldots \cup D_m$ *using a finer grid* Σ_2 *and obtain a finer resolution reconstruction* $\hat{I}_{D^{(2)}}^{\Sigma_2}$ *from the equation* $g = A_{D^{(2)}}^{\Sigma_2} \hat{I}_{D^{(2)}}^{\Sigma_2}$ *using the AC3DR algorithm.*

3. Further resolution refining: *If necessary, step 2 can be recursively applied until the object space is reconstructed at the desired resolution, which should not exceed the resolving capability of the optical system.*

Remark. Simulation results have shown that a suitable resolution choice is essential for obtaining reliable regularized solutions. Oversampling a region at a resolution which exceeds the resolution capability of the physical system leads to a more blurred reconstruction of the object, though correctly shaped and positioned. An illustrative example is presented in section 7.

The output of the hierarchical application of the adaptive constrained regularization is an *intensity map* of the object space. This map allows the segmentation of a multiple object scene into separate components. The segmentation is facilitated by the hierarchical structure of the algorithm.

6. Depth map construction using a weighted Durbin–Willshaw scheme

The final purpose of the present work is to draw the *depth map* of object space, i.e. the contour of the scene, from the intensity map provided

by the hierarchical adaptive constrained 3D-reconstruction algorithm. Therefore, for each horizontal plane $x = x_0$ an optimal curve, defined by an ordered set of points $w = (w_1, \ldots, w_m)$, which fits the high intensity reconstructed points is sought. A weighted Durbin-Willshaw ((Durbin and Willshaw, 1987),(Kung, 1993)) scheme has been used for this purpose. It consists in finding a set of points w which minimizes the following energy function:

$$E(w) = -\rho^2 \sum_{\mu} \log[r(p_\mu) \sum_{j} \exp(-|p_\mu - w_j|^2/(2\rho^2))] + \frac{\alpha}{2} \sum_{j} |w_{j+1} - w_j|^2 \,, \tag{11}$$

where $p_\mu \,, \mu = 1, \ldots, l$, are the high intensity points obtained at thresholding, ρ determines the effective attraction radius of each point p_μ, and α assures that the consecutive w_j are reasonably close to each other. $r(p_\mu)$ is a number proportional to the square of the intensity of the point p_μ.

The updating rule is the following:

$$\Delta w_i = \eta(\sum_{\mu} r(p_\mu)\Lambda^\mu(i)(p_\mu - w_i) + \alpha(w_{i+1} - 2w_i + w_{i-1})) \,, \tag{12}$$

where $\Lambda^\mu(i)$ is defined as:

$$\Lambda^\mu(i) = \frac{\exp(-|p_\mu - w_j|^2/(2\rho^2))}{\sum_j \exp(-|p_\mu - w_j|^2/(2\rho^2))} \,. \tag{13}$$

ALGORITHM 3 (Weighted Durbin–Willshaw scheme).

1. – *Consider the high intensity points* $p_\mu \,, \mu = 1, \ldots, l$ *and their associated intensities* $i_\mu \,, \mu = 1, \ldots, l$. *Compute the rating coefficients:* $r(p_\mu) = (i_\mu / \min_\mu(i_\mu))^2 \,, \mu = 1, \ldots, l$.
   – *Consider an arbitrary initial solution* $w^{(0)} = (w_1^{(0)}, \ldots, w_m^{(0)})$.

2. *Update the current solution according to the gradient descent rule:*

$$w^{(k+1)} = w^{(k)} - \eta \nabla w^{(k)} \,,$$

where $\eta > 0$ *and the gradient* $\nabla w^{(k)}$ *is given by:*

$$\nabla w_i^{(k)} = \sum r(p_\mu)\Lambda^\mu(i)(p_\mu - w_i^{(k)}) + \alpha(w_{i+1}^{(k)} - 2w_i^{(k)} + w_{i-1}^{(k)}) \,,$$

for any $i = 1, \ldots, m$.

Step 2 is iterated until convergence is reached.

In the case of multiple object scenes, the contour of each object can be drawn using a separate weighted Durbin–Willshaw structure.

This method of producing contours has the outstanding property of not being bound to the sampling points used at reconstruction. The points on the elastic Durbin-Willshaw net are floating in a *continuous object space*, not points of the sampling grid. In this way, the scheme overcomes inherent resolution limitations and problems posed by the interpolation at the sampling points, providing realistic contours of the objects in the scene.

7. Simulation results

The hierarchical adaptive constrained 3D-reconstruction algorithm has been tested on a set of computer generated images containing various object shapes and depths. The results obtained for images containing a 9 disparate point configuration, a two transparent thin face scene, and an opaque cube are presented in the following.

For the simplicity of the presentation, planar sections parallel to the Oyz-plane of the object space reconstruction are shown. The objects considered are all perpendicular to the Oyz-plane, so their intersections with $x = x_0$ planes are similar. Hence, recovering one section means recovering the whole object. Figures 5-16 contain both the object reconstruction and the original object for comparison. For graphic clarity purposes, the results are presented in reversed gray map (i.e. the darker gray shades represent higher levels of intensity).

1. *The hierarchical reconstruction of a 9 point scene*

   The image of a scene containing 9 disparate points has been processed.

   - *First hierarchical level: initial sampling and regularization.* A first regularization, performed on a wide sampling, has provided the reconstruction shown in Figure 5(a), whose 3 bin histogram rule thresholded variant is shown in Figure 5(b). The reconstruction is not completely accurate, therefore further local regularizations have been performed.
   - *Second hierarchical level - local regularization at the same resolution (algorithm 2 2a).* The first variant is to cut a rectangular region around the high intensity points (Figure 6(a)) and perform local regularization on it. The result and the

thresholded version are shown in Figures 6(a)-(b) and constitute a 100% accurate reconstruction. A more computationally economic alternative to the rectangular cut is to perform local regularization only on the region formed by the high intensity points and their neighbours (Figure 7(a)). The solution (Figures 7(b)-(c)) is again a 100% accurate reconstruction.

The scene contains disparate points, so the intensity map obtained by hierarchical regularization is also a depth map.

This is a 'toy' example which shows that the procedure works very well in the ideal case when the object points are among the sampling points. In the case of continuous objects, like the ones concerned in the following two examples, the solution is an interpolation of the real intensity values at the sampling points. This is why the Durbin–Willshaw scheme is necessary to decide on the final reconstructed contour of the object space.

2. *The hierarchical reconstruction of a two object scene*

   The object scene contains two transparent faces of unequal lengths perpendicular to each other. It is reconstructed using a course grid at the first hierarchical level and a finer grid on selected regions at the second one.

   – *First hierarchical level: initial sampling and regularization.* The reconstruction at low resolution is shown in Figure 8(a). The objects are correctly positioned and separated in the reconstruction, though their shapes are perturbed.

   – *Second hierarchical level: localized regularization with increased resolution (algorithm 2.2b).* Further processing is performed for the marked regions around the two objects and more accurate reconstructions are obtained (Figures 8(b)-(c)) using the *multiresolution* approach.

   – *Depth map drawing.* Finally, the contours of the objects are drawn using two independent Durbin–Willshaw structures. (Figures 9(a)-(b)).

3. *The hierarchical reconstruction of a transparent cylinder*

   – *First hierarchical level: initial sampling and regularization.* The processing of the image of a transparent thin cylinder yielded the scene depicted in Figure 10(a). The object recovered at this stage presents lateral (y) symmetry and is located

around the true position of the object, but the shape not clear. This is not unexpected, as the sampling grid is very sparse in the object space and cannot retain the curvature of the object's contour at this resolution.

- *Second hierarchical level – local regularization at the same resolution (algorithm 2 2a). Local regularization* is performed in the vicinity (dashed line in Figure 10(a)) of the high intensity points. The result is shown in Figure 10(b). The new reconstruction contains less high intensity points than the first one. It is noticed that it suggests a hollowed contour, the 'front' of the object (left hand side in the picture) being more prominent.
- *Depth map drawing.* The weighted Durbin–Willshaw scheme produced the contour depicted in Figure 11(a). It matches the actual shape of the object, shown in continuous line on the same figure. Figure 10(b) shows a 3D view of the reconstruction of the transparent cylinder.

4. *The hierarchical reconstruction of a an opaque cube*

- *First hierarchical level: initial sampling and regularization.* The processing of the image of the visible sides of an opaque cube yielded the scene depicted in Figure 12(a), which by 3 bin histogram rule thresholding gives the reconstruction shown in Figure 12(b). The object recovered at this stage is symmetric and the two sides are clearly separated.
- *Second hierarchical level – local regularization at the same resolution (algorithm 2.2a).* For better accuracy, *local regularization* is performed in the vicinity (Figure 13(a)) of the high intensity points. The result and its thresholding are shown in Figures 13(b)-(c). It can be noticed that the final reconstruction follows the actual contour of the object.
- *Depth map drawing.* In order to link the resulted 'edges', the weighted Durbin–Willshaw sheme described in section 6 has been used. The curve fitting the intensity map from Figure 13(c) is depicted in Figure 14. It matches very well the actual contour of the object, shown in continuous line on the same figure. A three dimensional view of the reconstructed object is presented in Figure 15.

A multiresolution regularisation (algorithm 2.2b) has also been tried for the opaque cube image, considering a denser sampling

grid (600 points instead of 136 points) in the area surrounding the high intensity region obtained at the first level reconstruction. The multiresolution regularization result is shown in Figure 16. It can be noticed that the reconstruction is more blurred than the one obtained previously by local regularization at the initial resolution. The blur is due to the fact that the new resolution for the object space exceeds the capability of the optical system.

8. Conclusions

The present paper has approached the object reconstruction and depth extraction from 3D-integral images as an inverse problem, which proves to be ill-conditioned. In order to cure the ill-conditioning of the problem, a hierarchical adaptive constrained 3D-reconstruction algorithm has been considered. It is based on the use of a sigmoid function to determine a bounded constrained least squares solution which provides information about the number of objects in the scene, their shape and absolute and relative position. The scheme can be hierarchically reapplied to the relevant regions of the object space in order to produce higher resolution and higher accuracy intensity maps of the scene. Finally, a weighted Durbin–Willshaw scheme is employed to produce a depth map of the object space from the intensity map. The simulation results have shown that the number of objects, absolute and relative positions and shape are recovered accurately. The hierarchical scheme produces high resolution reconstructions and is computationally efficient, as only the relevant regions of the scene are oversampled.

References

G. Lippmann. La Photographie integrale. *Comtes Rendus*, Academie des Sciences, 146:446–451, 1908.

T. Okoshi. *Three dimensional imaging techniques.* Academic Press, 1976.

N. Davies, M. McCormick. Holoscopic imaging with true 3D-content in full natural colour. *J.Phot.Science*, 40:46–49, 1992.

N. Davies, M. McCormick, M. Brewin. Design and analysis of an image transfer system using microlens arrays. *Opt. Eng.*, 33(11):3624–3633, 1994.

S. Manolache, A. Aggoun, M. McCormick, N. Davies. A mathematical model of a 3D-lenticular integral recording system. *Proceedings of IEEE Vision, Modeling and Visualization*, 51–58, Erlangen, 1999.

M. Bertero, P. Boccacci. *Introduction to inverse problems in imaging.* Institute of Physics Publishing, 1998.

A.N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4:1035–1038, 1963.

J.M. Ortega, W.C. Rheinboldt. *Iterative solutions of nonlinear equations in several variables.* Academic Press, 1970.

R. Durbin, D.J. Willshaw An analogue approach to the travelling salesman problem using an elastic net method. *Nature*, 326, 16:689, 1987.

S.Y. Kung. *Digital neural networks.* Prentice-Hall, 1993.

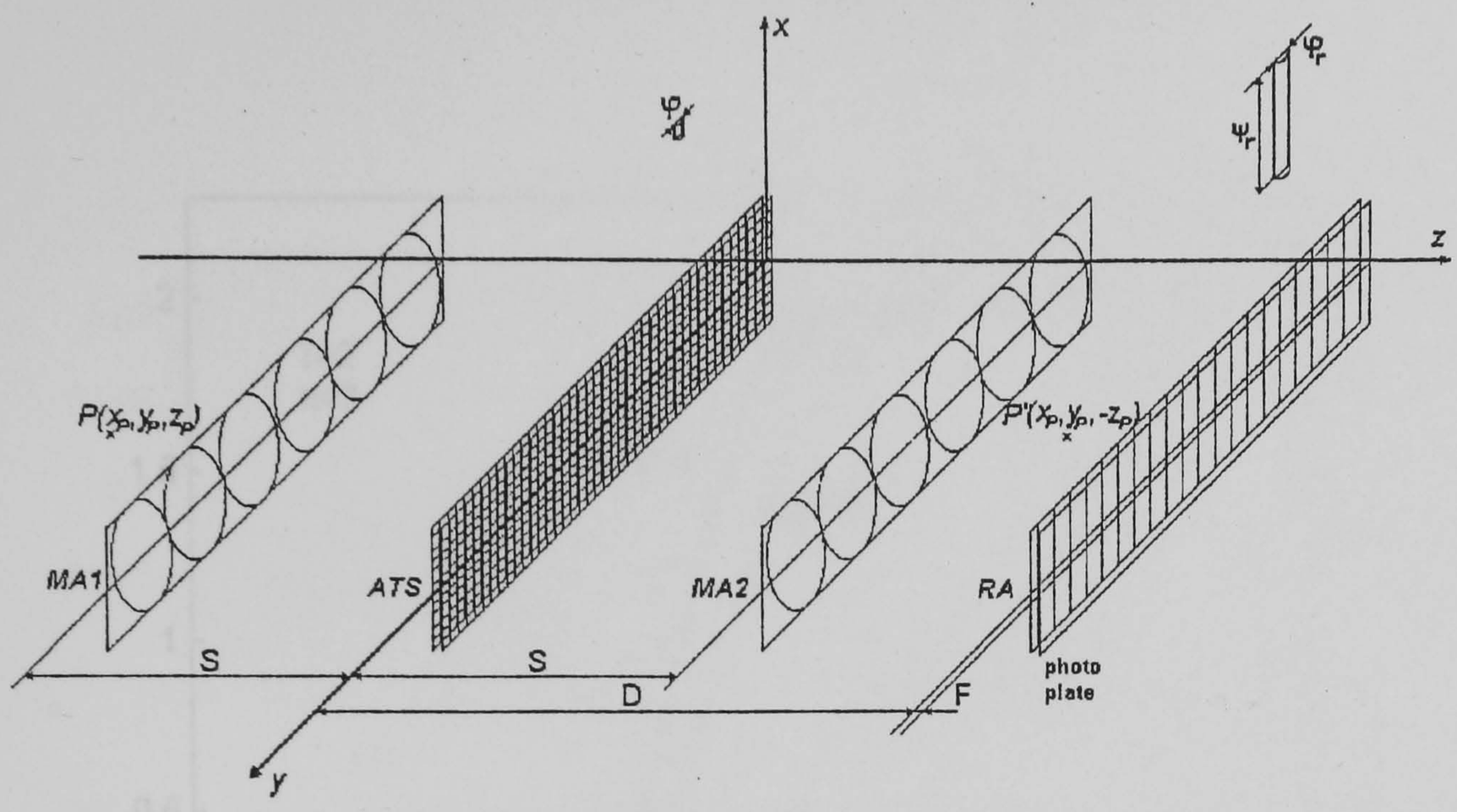


Figure 1. 3D-lenticular integral imaging camera system

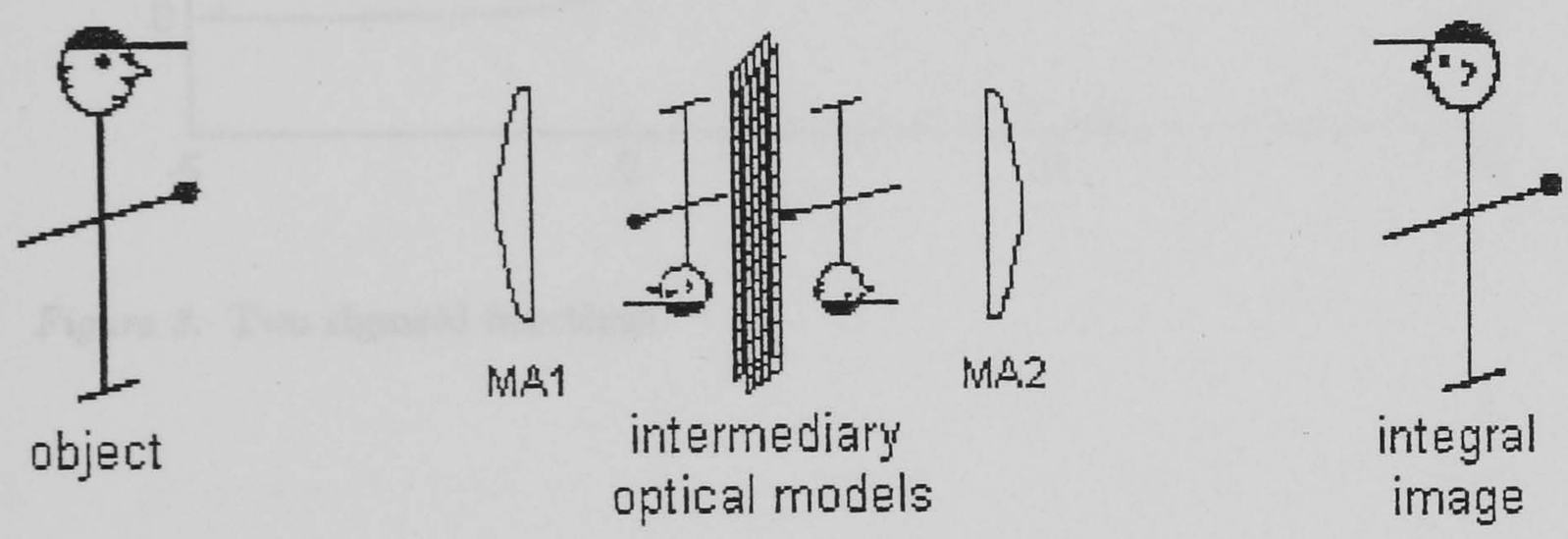


Figure 2. Optical stages in the integral imaging

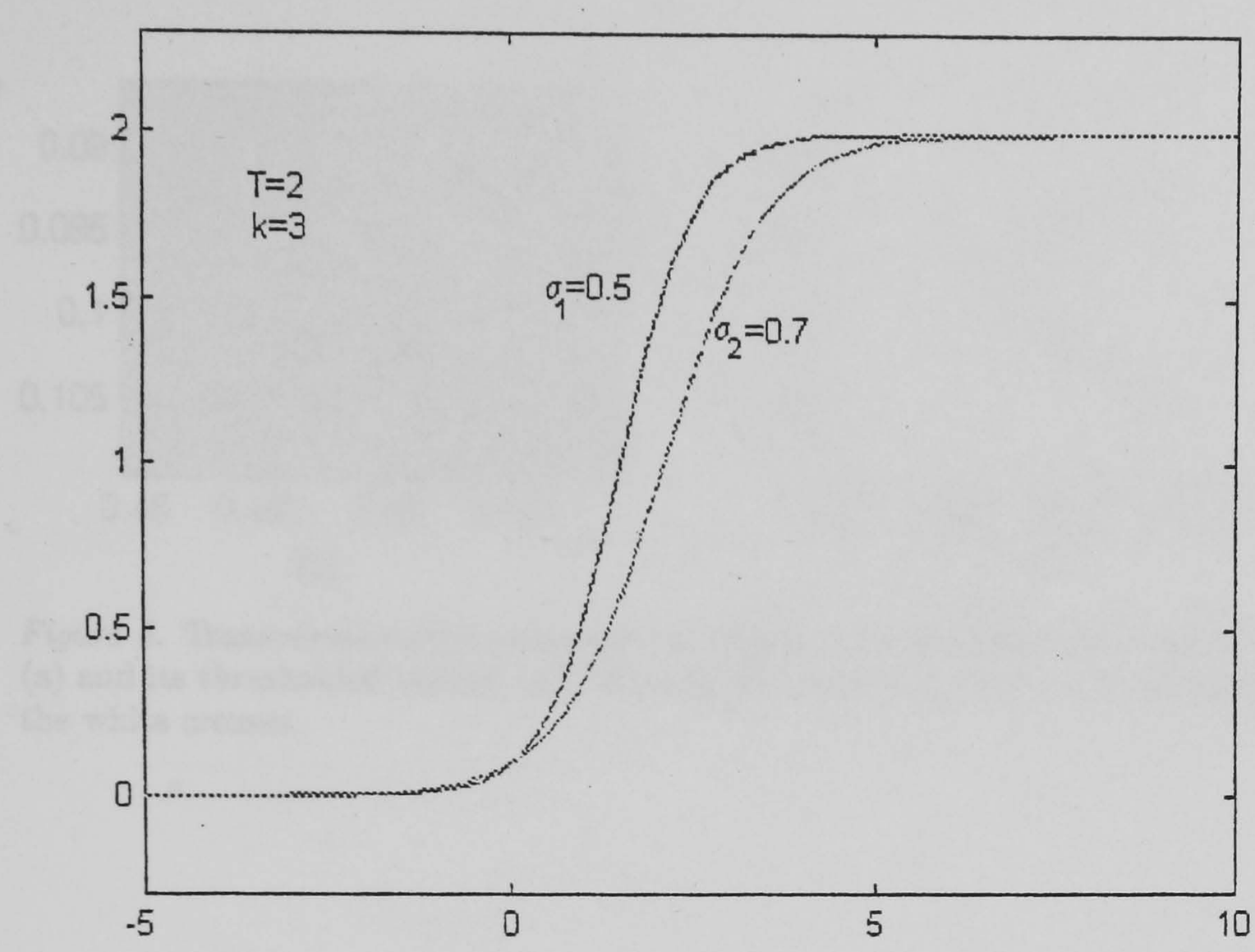


Figure 3. Two sigmoid functions.

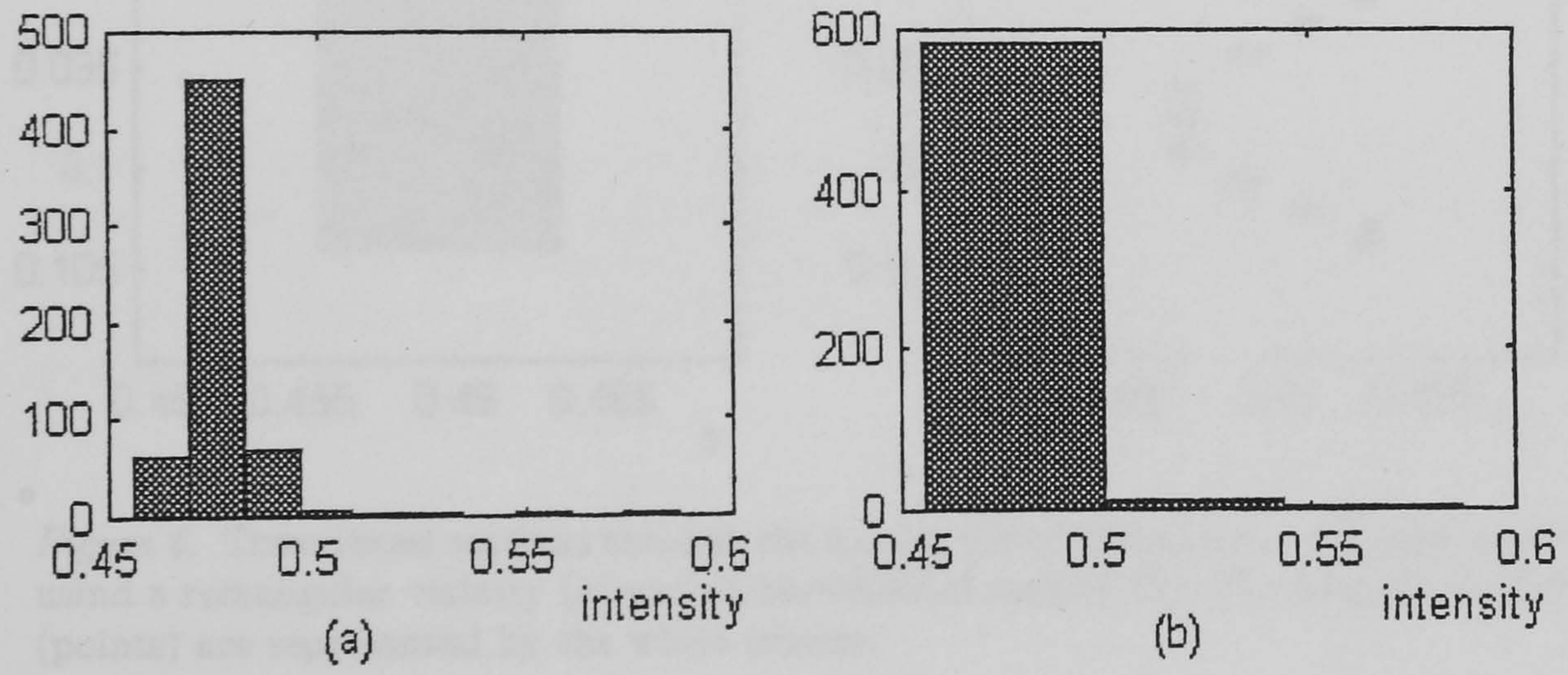


Figure 4. A 10-equidistant bin and a 3-equidistant bin histogram of a reconstructed intensity.

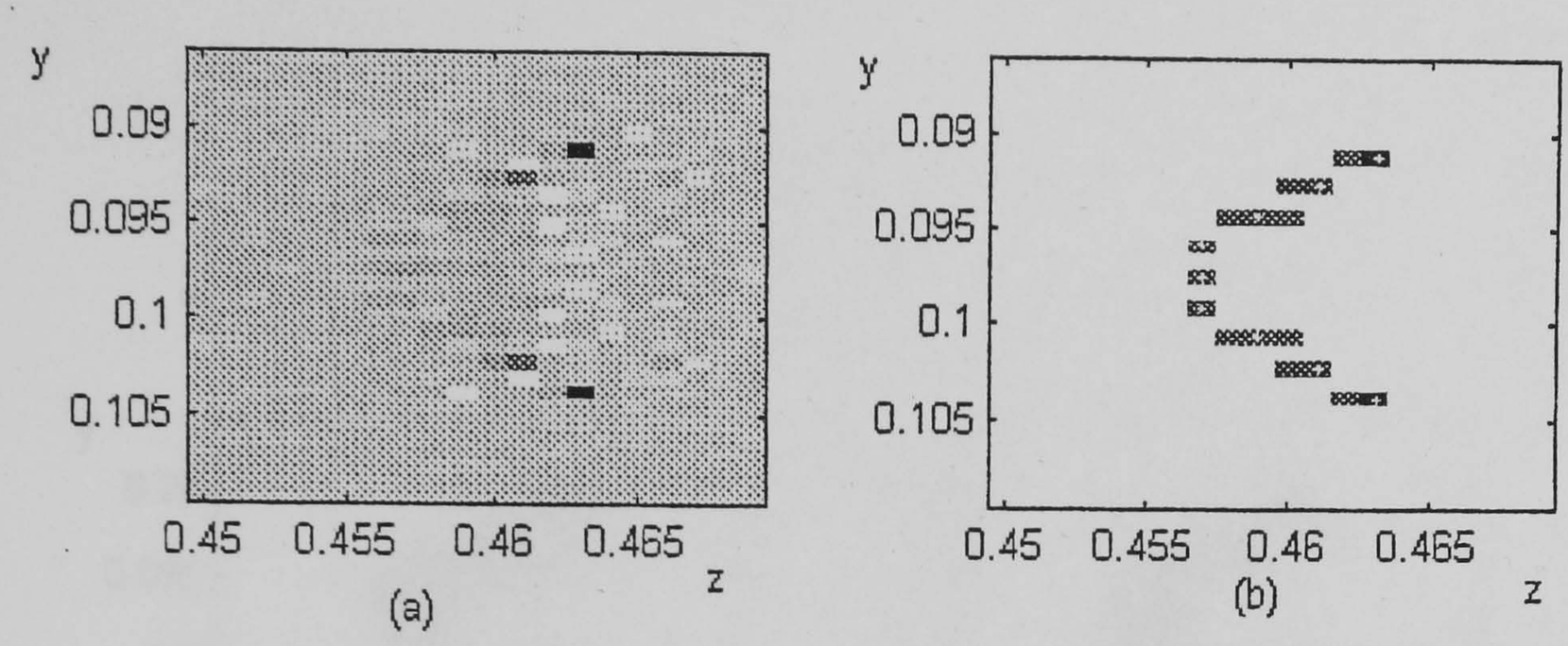


Figure 5. Transversal sections through the initial reconstruction of a 9 point scene (a) and its thresholded version (b). The original objects (points) are represented by the white crosses.

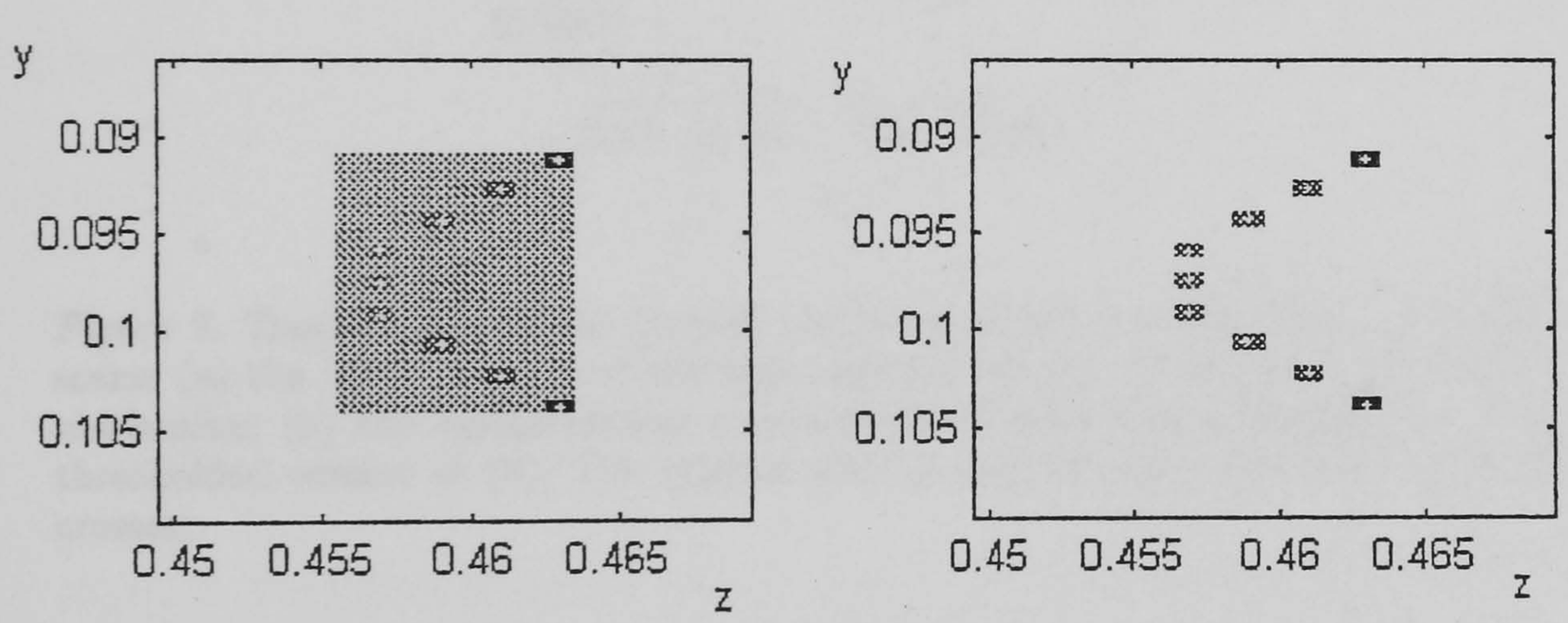


Figure 6. Transversal sections through the localized reconstruction of a 9 point scene usind a rectangular vicinity (a) and its thresholded version (b). The original objects (points) are represented by the white crosses.

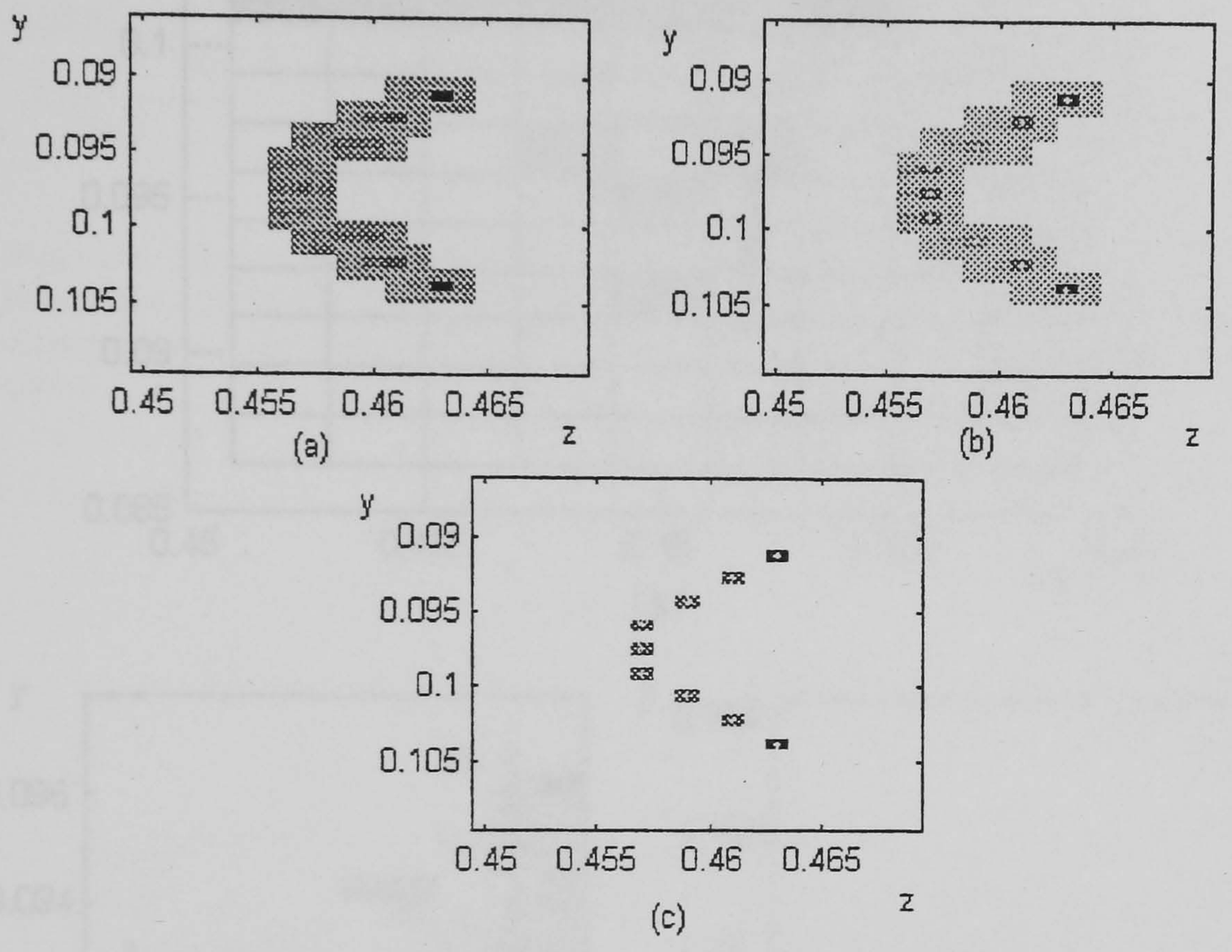


Figure 7. Transversal sections through the hierarchical reconstruction of a 9 point scene: (a) the 'close' vicinity of the high intensity points selected for localized regularization; (b) the reconstruction obtained after localized reconstruction; (c) the thresholded version of (b). The original objects (points) are represented by white crosses.

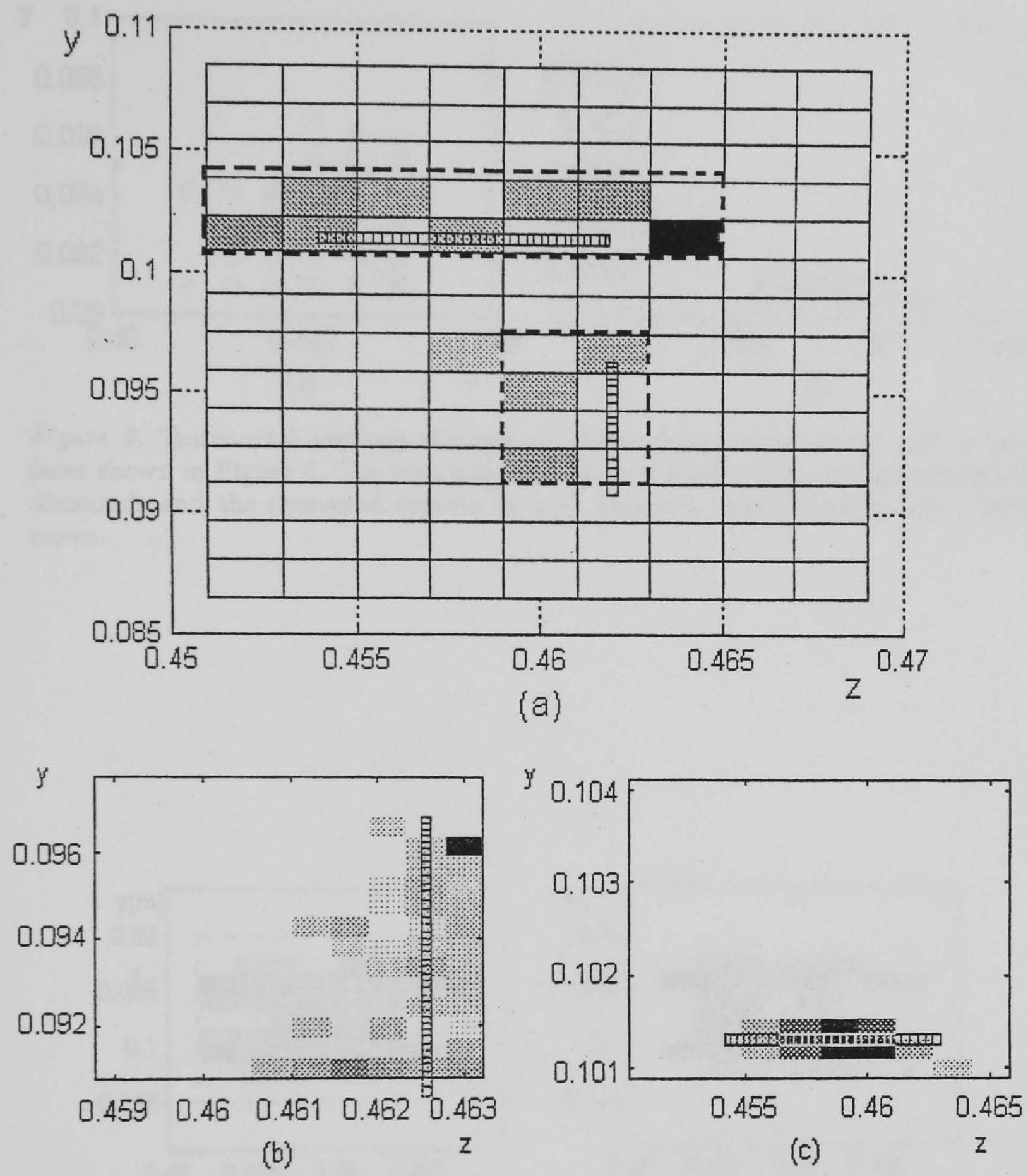


Figure 8. Transversal sections through the hierarchical reconstructions of a pair of thin faces, one orthogonal and another parallel to the recording screen: (a) reconstruction at course sampling; (b)-(c) reconstructions after localized regularization at increased resolution aroud the two objects. The original objects are drawn in continuous line.

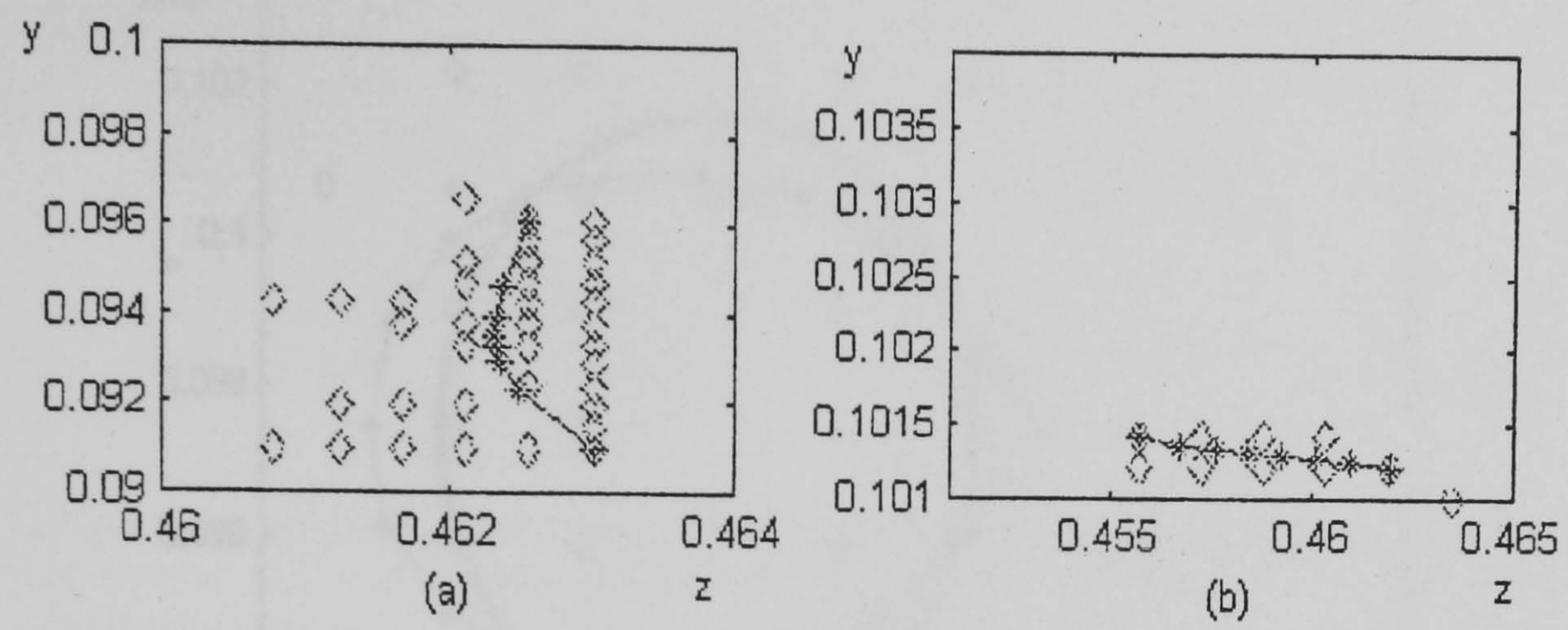


Figure 9. Transversal sections through the local depth maps of the pair of thin faces shown in Figure 8. The high intensity reconstructed points are represented by diamonds and the recovered contour of each object is represented by the starred curve.

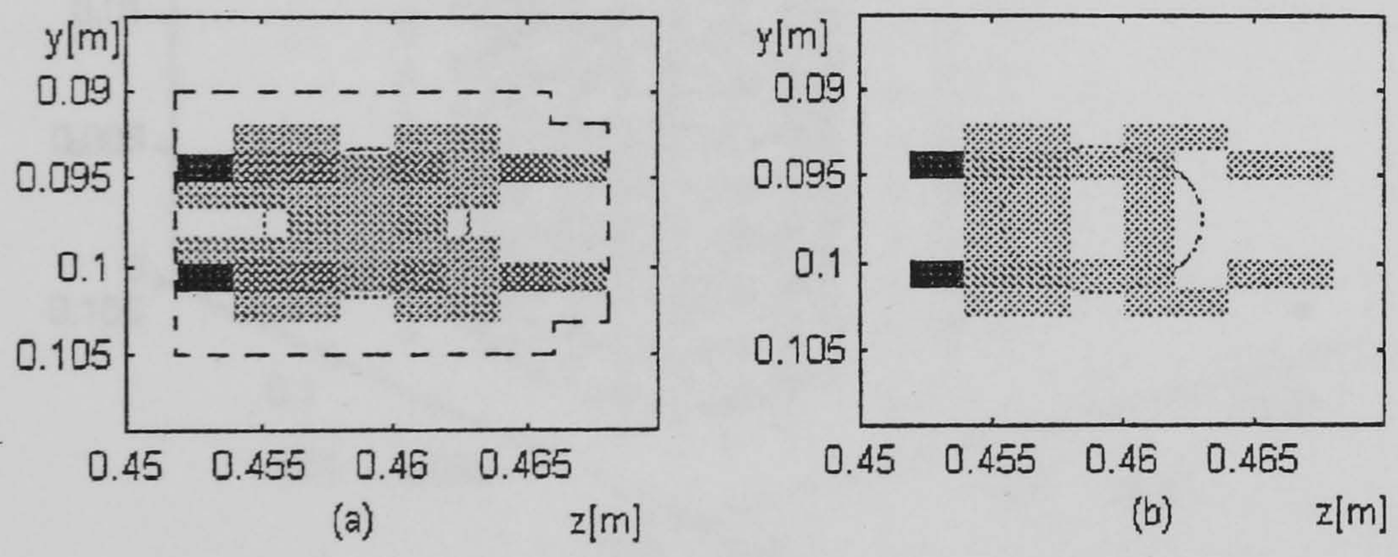
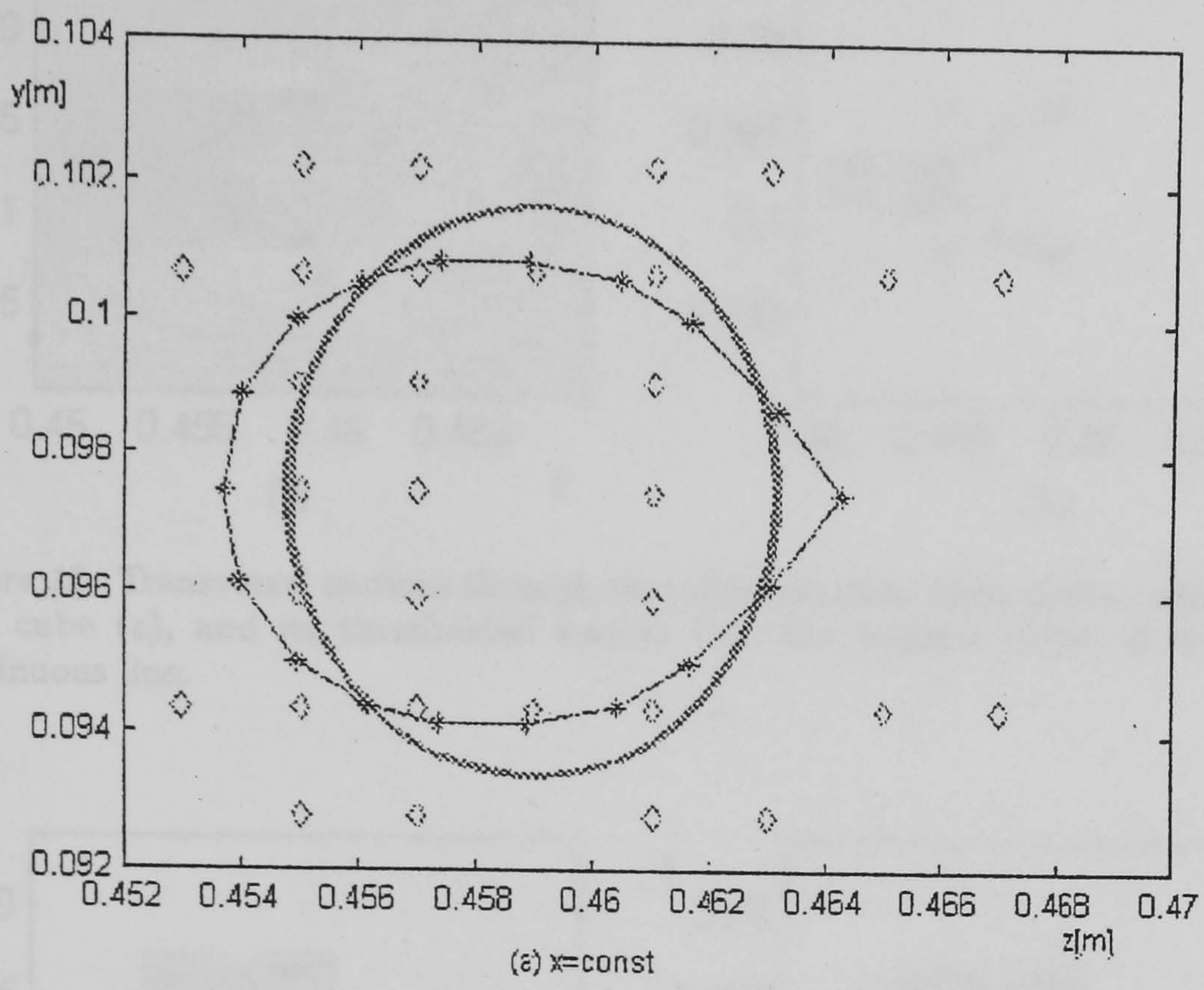


Figure 10. Hierarchical AC3D reconstruction of a transparent cylinder from a noisy integral image: (a) initial regularization and high intensity region selection (dashed line); (b) localised regularization at the same resolution in the region selected at (a). The original object is figured in continuous line.



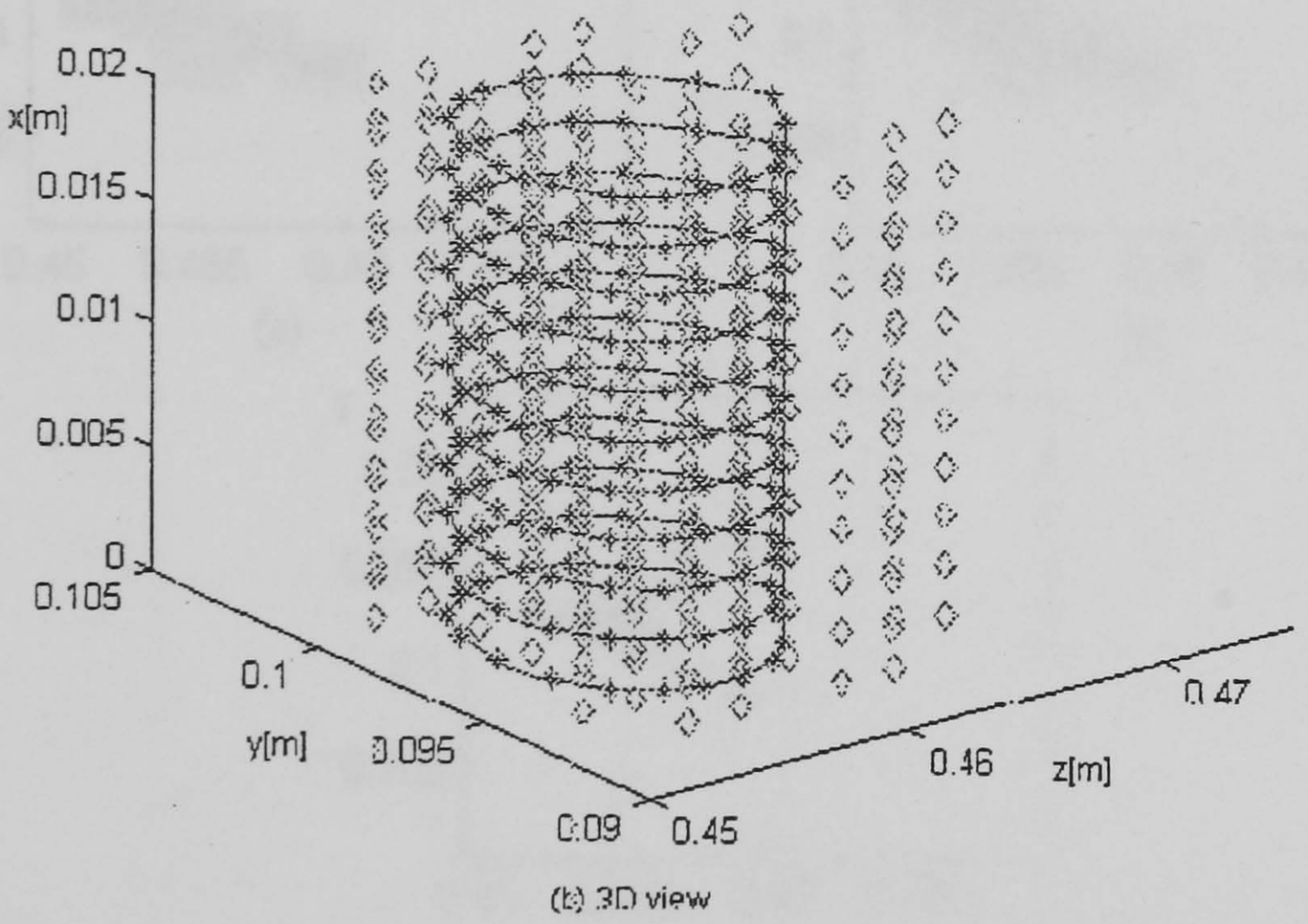


Figure 11. Durbin-Willshaw depth map corresponding to the HAC3D reconstruction of a transparent cylinder, depicted in Figure 10: (a) Transversal section with a plane $x = const$. The original object is figured in continuous line. (b) 3D view of the reconstruction.

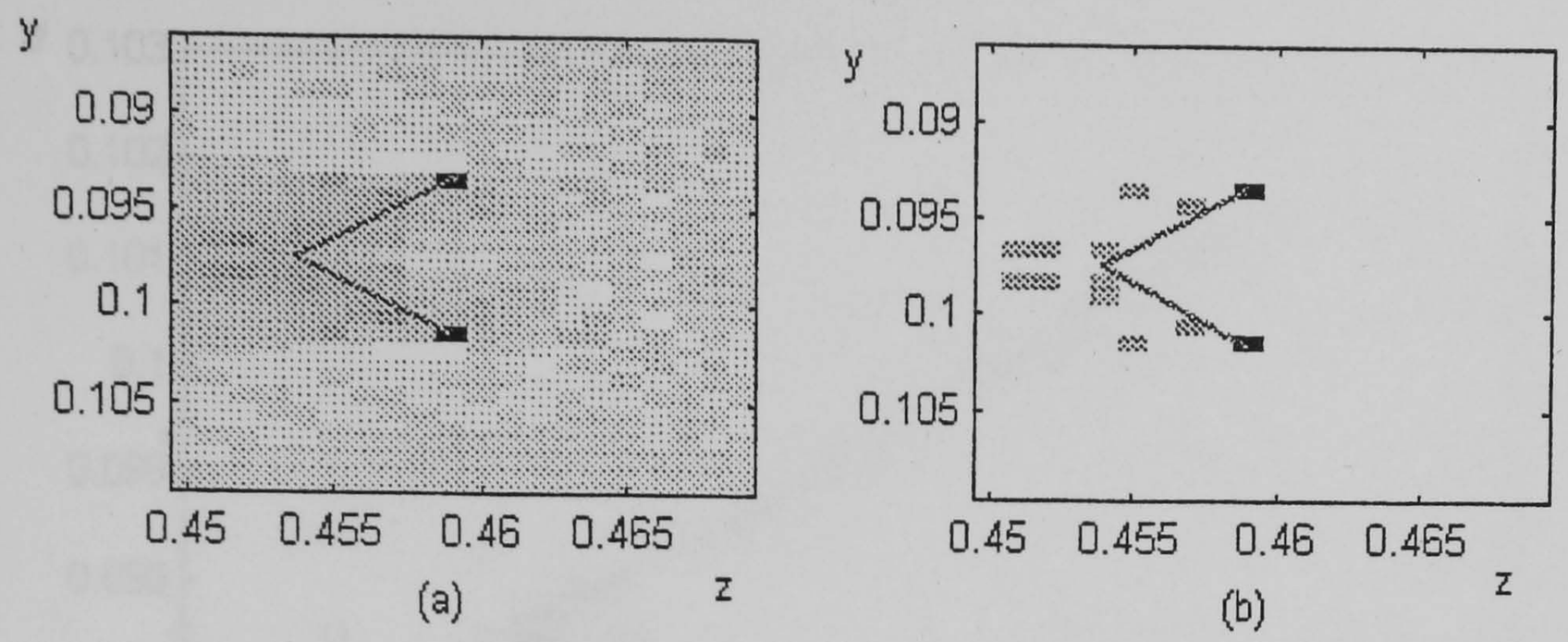


Figure 12. Transversal sections through the initial reconstruction of the visible sides of a cube (a), and its thresholded version (b). The original object is drawn in continuous line.

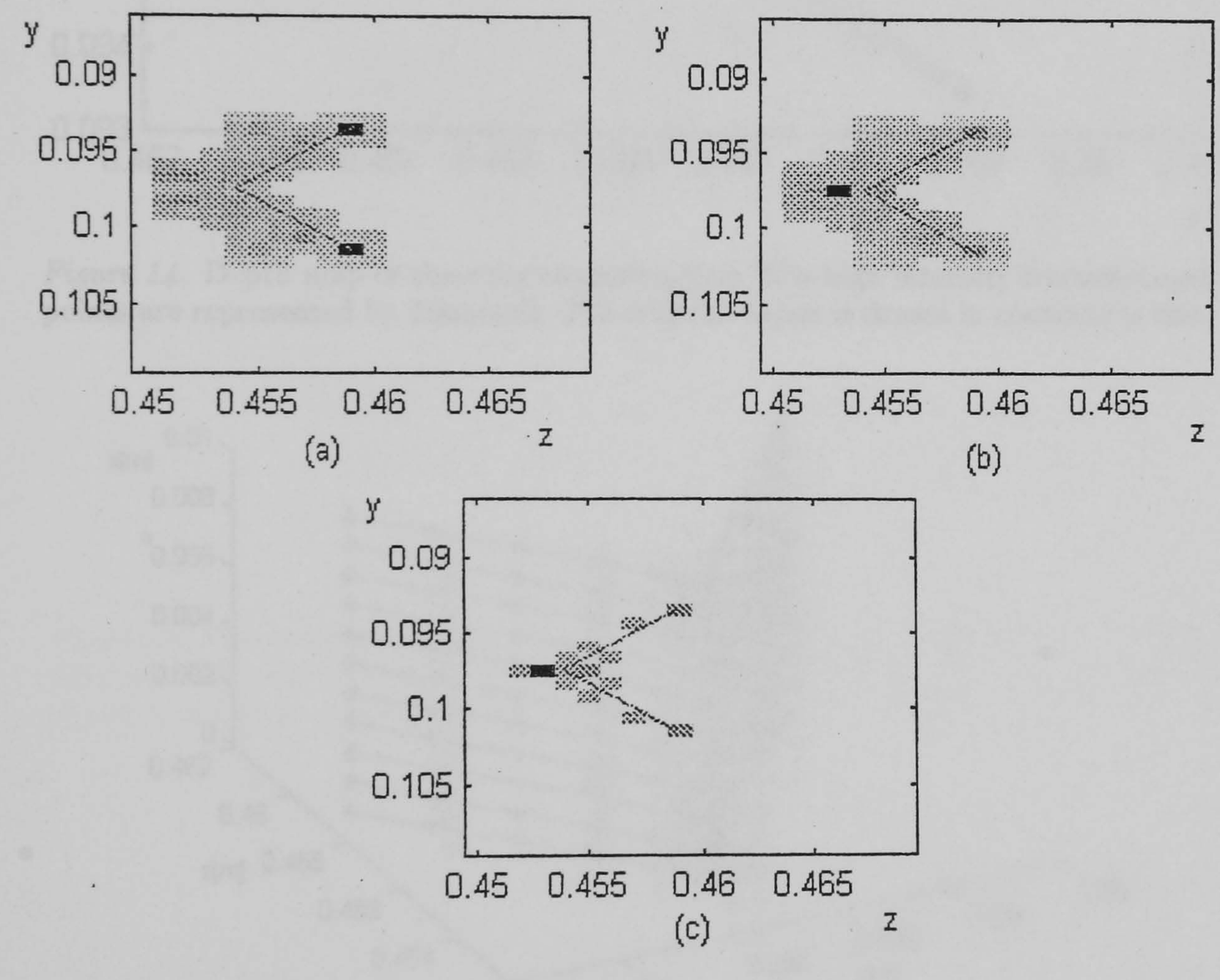


Figure 13. Transversal sections through the localized reconstructions of the visible sides of a cube: (a) selected vicinity of the high intensity points reconstructed at the first regularization; (b) reconstruction after localized regularization at the same resolution; (c) thresholded result. The original object is drawn in continuous line.

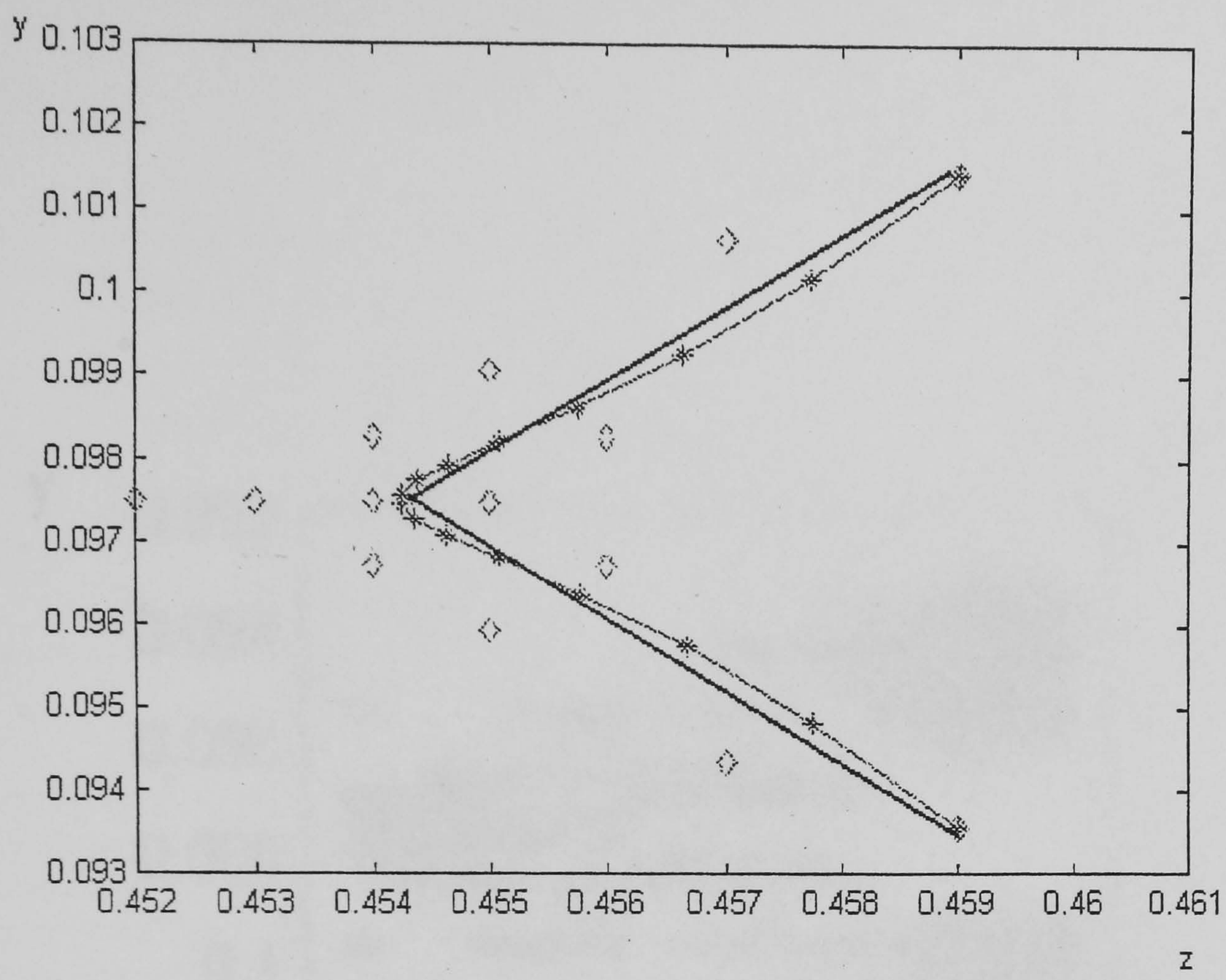


Figure 14. Depth map of the cube reconstruction. The high intensity reconstructed points are represented by diamonds. The original object is drawn in continuous line.

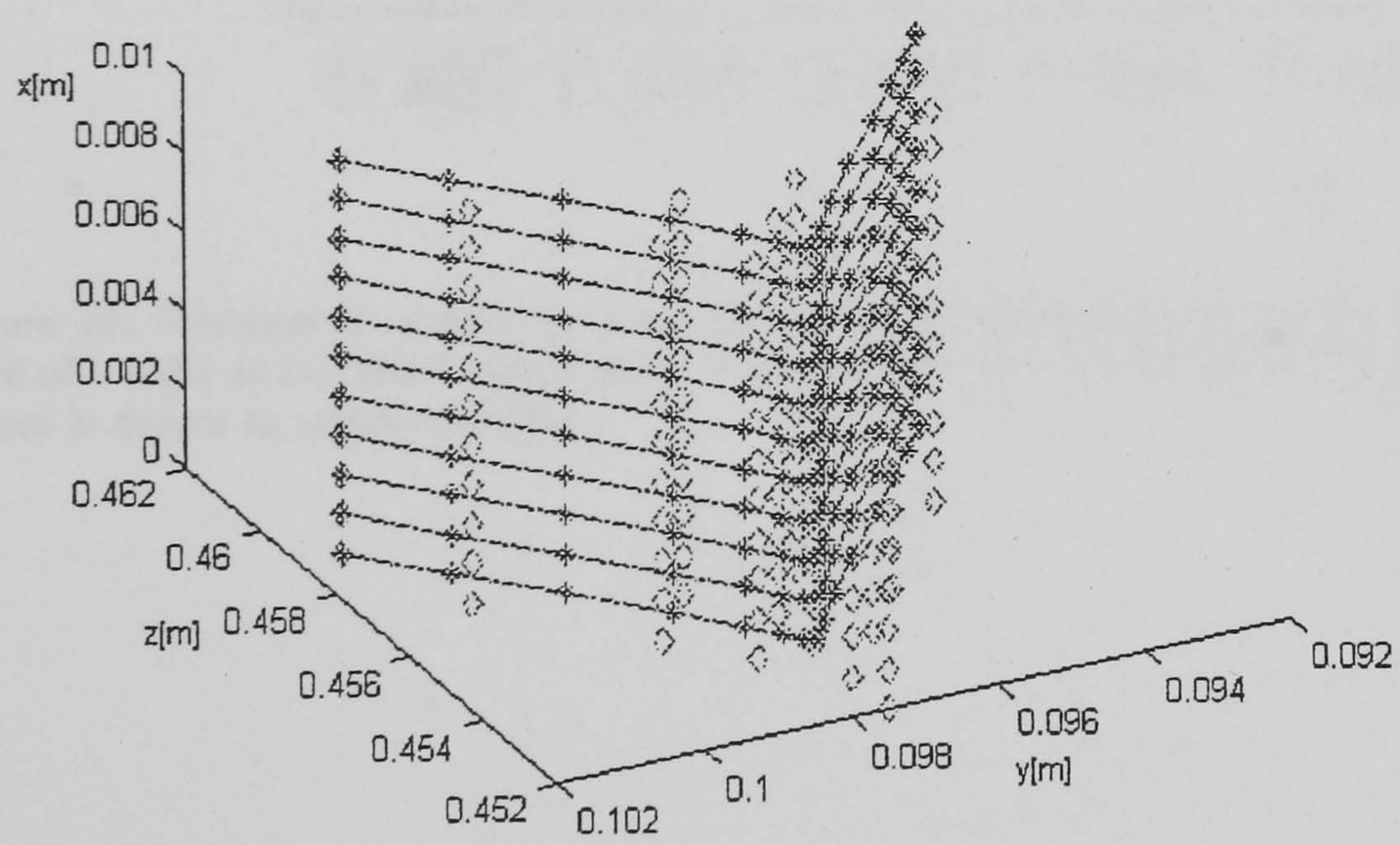


Figure 15. Durbin-Willshaw depth map corresponding to the HAC3D reconstruction of the visible faces of an opaque cube depicted in Figure ?? (3D view).

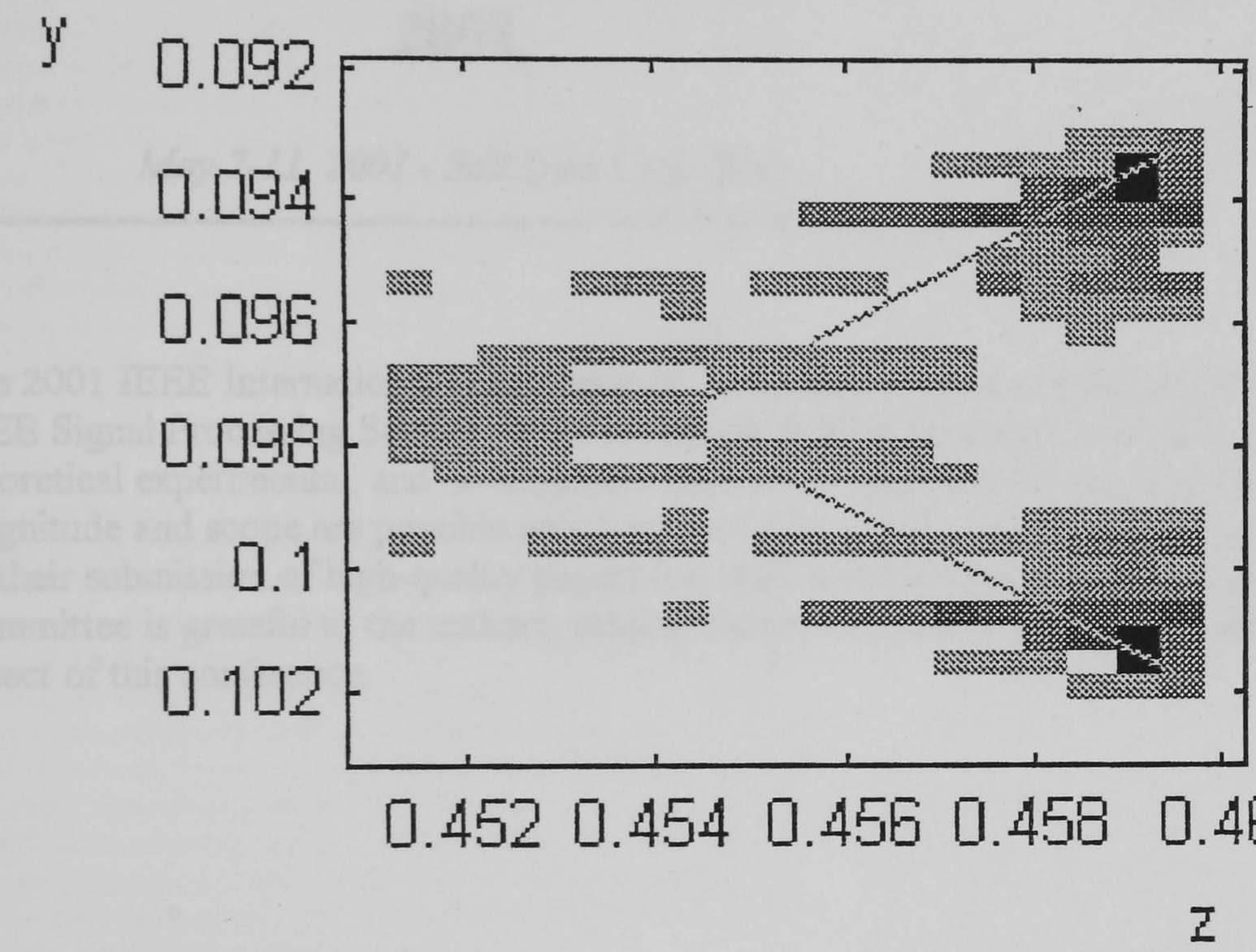
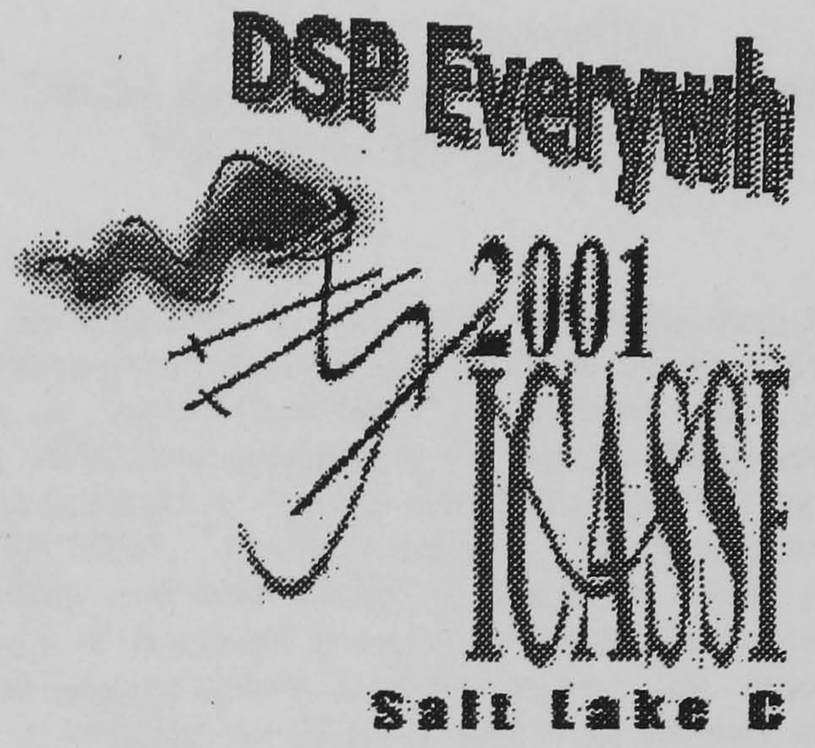


Figure 16. Transversal section through the localized reconstruction of the visible sides of a cube at increased resolution. The blur effect can be noticed. The original object is drawn in continuous line.

IEEE Signal Processing Society

International Conference on Acoustics, Speech, and Signal Processing 2001

May 7-11, 2001 - Salt Lake City, Utah



The 2001 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2001) IEEE Signal Processing Society, is the twenty-sixth in an ongoing series of annual international confer theoretical experimental, and development aspects of signal processing, acoustics, and speech. Confere magnitude and scope are possible only because of the continuing interest and support of Society memb by their submission of high-quality papers and their participation in the conference. The ICASPP 2001 Committee is grateful to the authors, session chairs, volunteers, and all the other people who have cont aspect of this conference.

HIERARCHICAL ADAPTIVE REGULARISATION METHOD FOR DEPTH EXTRACTION FROM PLANAR RECORDING OF 3D-INTEGRAL IMAGES

Silvia Manolache, Malcolm McCormick*

De Montfort University
Department of Engineering and Technology
The Gateway, Leicester LE1 9BH, UK

S.-Y. Kung

Princeton University
Department of Electrical Engineering
Princeton, NJ-08544, USA

ABSTRACT

The paper presents a novel algorithm for object space reconstruction from the planar (2D) recorded data set of a 3D-integral image. The integral imaging system is described and the associated point spread function is given. The space data extraction is formulated as an inverse problem, which proves ill-conditioned, and tackled by using a hierarchical multiresolution strategy and imposing additional conditions to the sought solution. The hierarchisation strategy and the two-phase adaptive constrained 3D-reconstruction algorithm based on the use of two sigmoid functions are presented. Finally, illustrative simulation results are given.

1. INTRODUCTION

The development of 3D-imaging systems has been a constant pursuit of the scientific as well as of the entertainment community in the new technological era. Integral photography was pioneered by Lippmann ([1], 1908), who used microlens arrays to create, record on photographic film, and replay integral three dimensional images. Since 1908, when it was first reported, the integral photographic technique has been improved as a result of theoretical studies, technical innovations of the optical systems, and progress in microlens manufacturing. Integral imaging resembles to holography, but it uses natural light and reproduces true colour optical models. Hence, it offers a viable alternative to other autostereoscopic systems ([2]).

The optics of an advanced form of integral imaging system employing a two tier optical network in which a true 3D optical reconstruction of a scene is transferred to the capture device, has been described in detail by Davies and McCormick ([3], [4]).

Extracting depth information from 3D-integral images has various applications, which range from remote inspection in robotic vision and medical imaging to combining real and computer generated 3D-pictures in a virtual studio. The aim of the present paper is to describe a method of reconstructing the composition of a 3D-object space from 2D-recorded data encoding the scene. The algorithm is applicable both to integral imaging and holography. The current work has dwelt upon the application to integral imaging. More specifically, the depth estimation from 3D-integral pictures is formulated as an inverse problem: given the image, i.e. the recorded 'effect' of the object space, find the 'cause' which had produced it, i.e. recover the intensity contours and, consequently, the composition and the depth of the object scene. The direct problem - image formation and recording - has been studied in a previous paper, and the equation of the point spread function for the lenticular integral imaging system has been derived ([5]). Inverse problems in imaging are ill-posed and their discrete correspondents are ill-conditioned due to the inherent loss of information associated with the direct process ([6]). In order to cure the ill-posedness of the problem, approximate solutions satisfying additional constraints coming from the physics of the problem are searched. The present work comprises an adaptive regularisation scheme for obtaining a constrained least squares solution of the depth extraction problem, which is hierarchically applied in order to obtain high resolution object space reconstruction.

2. 3D-INTEGRAL IMAGING SYSTEM AND ASSOCIATED POINT SPREAD FUNCTION

The optical arrangement of the 3D-integral imaging system, shown in Figure 1, comprises two macrolens arrays (MA1 and MA2) placed equidistantly behind and in front of an autocollimating transmission screen (ATS). The ATS is made up of two microlens arrays separated by their joint focal distance. The recording plane, a photographic plate whose position coincides with the focal plane of a microlens array (RA), lies within the optical model.

The paper considers a camera configuration where the ATS contains square based hemispherical lenslets and the recording array is made up of identical semicylindrical microlenses. This system is known as a *3D-lenticular integral imaging system* ([3], [5]) and produces 3D-images containing horizontal parallax.

The object is imaged by the input macrolenses which transmit compressed transposed images that occur at or near the central double microlens screen (ATS). The screen inverts the optical sense of each intermediary image (see Figure 2), and, simultaneously, presents these spatially reversed 3D-optical models to the corresponding output macrolenses. The output macrolenses retranspose the optical models to the correct spatial location. The final integrated image, formed by superposed optical models projected by the second macrolens array, is a true 3D optical 1:1 recon-

*The first author performed the work while at Princeton University under the sponsorship of the Exchange Visitor Program P1 180 between Princeton University and De Montfort University.

struction of the object. It is recorded as a *2D-sampled data set* possessing either horizontal parallax (if semicylindrical microlenses are used in the recording array), or continuous parallax (in the case of the spherical or square based microlens arrangement).

The integral image recorded in the focal plane of a recording microlens array as a *planar* sampled data set contains all the 3D information related to the object space. Each microlens of the recording array samples a fractional part of the scene, many microlenses recording directional information of the scene from different viewing angles. Therefore, parallax information about any particular point is spread over the recording plane. Redisplay of the full spatial model as a real 3D image can be effected by overlaying the sampled data set by an integral decoding element.

The spread function of the entire optical process combines the spread effect of both the central double microlens screen and the recording array ([5]):

$$K(X,Y,x,y,z) = \sum_{l(x,y,z)} K_{l(x,y,z)}\,, \qquad (1)$$

where (X,Y) are the coordinates in the image space, (x,y,z) represent the object space, and the summation is done with respect to the index of the microlenses which 'see' the point (x,y,z). In the above equation, $K_{l(x,y,z)}$ is the point spread function component behind one microlens of the recording screen:

$$K_{l(x,y,z)} = \alpha(z)\exp\left(-\frac{(X-X_l)^2}{a_1^2}\right)\exp\left(-\frac{(Y-Y_l)^2}{a_2^2}\right) \qquad (2)$$

The point (X_l,Y_l) represents the point of maximum intensity in the microimage formed behind microlens l and:

$$\begin{aligned}\alpha(z) &= \frac{v^4 w_1^2 w_2^2}{\sqrt{(v^2+w_1^2)(v^2+w_2^2)}} \\ a_1^2 &= v^2+w_1^2\,, \qquad a_2^2 = v^2+w_2^2\end{aligned} \qquad (3)$$

depend explicitly on lenslet parameters and point depth z ([5]).

Equation (2) indicates that the spread function is space variant and that it holds depth information in all the factors. This fact will be used in recovering three dimensional data about the object space. Due to the small values of the parameters involved in the point spread function formulae, only the significant factors should be considered in a computational approach of the depth extraction problem. When using lenticular arrays (i.e. semicylindrical microlenses), the spread on the x-dimension decreases rapidly to 0, so the x-factors in the point spread expression can be dropped.

3. FORMULATION OF THE OBJECT SPACE RECONSTRUCTION AS AN INVERSE PROBLEM

The image formation equation for the integral imaging system is:

$$g(X,Y) = \int I(x,y,z)\sum_{l(x,y,z)} K_l(X,Y,x,y,z)\,dx\,dy\,dz\,, \qquad (4)$$

where $I(x,y,z)$ is the intensity at point (x,y,z).

It is noticed that the integral image formation process is not a convolution process because the critical shift invariance property does not hold. Therefore classical deconvolution methods cannot be used to tackle the problem of scene reconstruction from an integral image.

However, the point spread function provides a linear image formation operator A such that equation (4) be written as:

$$g = A\,I\,. \qquad (5)$$

The reconstruction of the intensity distribution in the object is the inverse process of the image formation given by equation (5). The discrete variant of the problem has been treated considering various samplings of the object space. Numerical experiments have shown that the matrix associated with the point spread function is very ill-conditioned. Alternative methods of reconstructing the object space are thus necessary and one of them is presented in the next section.

4. HIERARCHICAL ADAPTIVE CONSTRAINED 3D-RECONSTRUCTION ALGORITHM

The ill-conditioning of the problem is determined by the existence of very small singular values of the operator A, which virtually increase the dimension of the null-space of A in numerical applications and make the solutions of the equation (5) be very unstable. The number of small singular values can be decreased by reducing the size of the matrix corresponding to A, so by considering only a low number of sampling points in the object space (lower than the number of pixels in the image). This operation has the drawback of inducing low resolution in the object reconstruction.

4.1. Hierarchical multiresolution strategy

The conflict between ill-conditioning of the imaging operator and high resolution reconstruction requirement has been solved by adopting a *hierarchical multiresolution strategy* based on zooming in the high interest regions. This approach consists of the following steps:

1. Coarse sampling: Sample a large domain $D^{(1)}$ using a coarse grid and obtain a low resolution reconstruction $\hat{I}_{D^{(1)}}$ from the equation $g = A_{D^{(1)}}\hat{I}_{D^{(1)}}$.

2. Fine sampling: Select the high interest regions $D_1,\ldots,D_m$ of the object space from the low resolution reconstruction $\hat{I}_{D^{(1)}}$ and consider the rest of the space as determined. Sub-sample the union $D^{(2)} = D_1 \cup \ldots \cup D_m$ and obtain a finer resolution reconstruction $\hat{I}_{D^{(2)}}$ from the equation $g = A_{D^{(2)}}\hat{I}_{D^{(2)}}$.

If necessary, the procedure can be recursively reapplied until the object space is reconstructed at the desired resolution.

The algorithm for obtaining a reliable solution for each sampled domain is given in the next subsection.

4.2. Adaptive constrained 3D-reconstruction algorithm

Choosing a low number of sampling points does not solve entirely the ill-conditioning problem, as the linear system (5) is still large, so unstable. Therefore, additional constraints coming from a priori knowledge about the object space have to be imposed upon the sought solution $\hat{I}$. More

precisely, the solution $\hat{I}$ has to be such that $A\,\hat{I}$ is as close as possible to the image g, i.e.:

$$||A\,\hat{I} - g|| = minimum\,. \qquad (6)$$

Also, a condition of positivity and bounding is necessary, as there are neither negative nor infinite intensity objects:

$$\hat{I} \in [0, T]\,. \qquad (7)$$

$\hat{I}$ satisfying (7) can be written as the thresholded value of arbitrary data u: $\hat{I} = f(u)$, where f is a thresholding function. The projection of the interval $[0, T]$ onto R is not differentiable, thus a sigmoid approximation of it, $f_\sigma : R^n \rightarrow [0, T]^n$, has been preferred:

$$f_\sigma(u_i) = \frac{T}{1 + exp\left(-\left(\frac{u_i}{\sigma} - k\right)\right)}\,,\ i = 1 \ldots n\,, \qquad (8)$$

where n is the number of components of the vector $\hat{I}$, i.e. the number of sampling points chosen in the object space. σ is an arbitrary number which controls the slope of the thresholding curve and k is a translation term ([7]).

To sum up, the sought vector $\hat{I}$ has the form $\hat{I} = f_\sigma(u)$ and minimises the discrepancy functional $\epsilon = ||A\,I - g||$.

The gradient of ϵ with respect to σ is:

$$\nabla_\sigma \epsilon = (A\,f_\sigma(u) - g)^T\,A\;diag(f'(u_i))_{i=1\ldots n}\,. \qquad (9)$$

The variation of u which leads to a constrained least-squares solution is on the decrease direction of the gradient $\nabla_\sigma \epsilon$:

$$\Delta u = -\eta \nabla_\sigma \epsilon\,, \eta > 0\,. \qquad (10)$$

The algorithm starts with an arbitrary value for u, e.g. $u = 0$, and iteratively modifies u with a quantity Δu computed as in the above equation.

Choosing $\sigma = T/k$, the function f_σ becomes a quasi-linear thresholding function. Its use in simulations has proved inefficient. Instead, a two-phase scheme has been preferred:

1. Use a relatively small value of σ, e.g. $\sigma = T/(12k)$, and perform a number of iterations.

2. Use a larger σ, e.g. $\sigma = T/(6k)$, until convergence.

The first phase of the algorithm performs a polarisation of the reconstructed intensity values, whilst the second phase removes arbitrary equalisations of the reconstructed values. The final regularised solution $f_\sigma(u)$ is thresholded and only the highest values are kept. This is the resulting object space reconstruction output by the current algorithm.

5. SIMULATION RESULTS

The hierarchical adaptive constrained 3D-reconstruction algorithm has been tested on a set of computer generated images. These images contain either a single face with various inclination angles with respect to the Oxy-plane, or two faces having equal or different inclination angles, or a cube. For the simplicity of the presentation, planar sections parallel to the Oyz-plane of the object space reconstruction are shown. The objects considered are all perpendicular to the Oyz-plane, so their intersections with $x = x_0$ planes are similar. Hence, recovering one section means recovering the whole object. Figures 3-5 contain both the object reconstruction and the original object for comparison. The darker gray shades represent higher levels of intensity in these representations.

Figure 3 depicts the object reconstruction from an image containing a single transparent thin face, 0.4 mm wide, 8 mm long, centred at the point (y=0.975 m, z=0.459 m), perpendicular to y-axis using the adaptive constrained reconstruction algorithm. The position and dimension of the object are correctly recovered even at low resolution.

The processing of the image of the visible sides of an opaque cube yielded the scene depicted in Figure 4(a). The object recovered at the coarse sampling stage is symmetric and correctly positioned in space, and the two sides are clearly separated. The region where the cube was detected, marked by the dashed line in Figure 4(a), has been further sampled and processed. The result is shown in Figure 4(b) and provides a more accurately contoured object.

The hierarchical reconstruction of a two object scene is presented in Figure 5. The two objects are two transparent faces of unequal lengths perpendicular to each other. They are correctly positioned and separated in the reconstruction, though their shapes are slightly perturbed. A further sampling and processing is performed for the marked regions around the two objects and more accurate reconstructions are obtained (Figures 5(b)-(c)).

The adaptive constrained reconstruction algorithm has converged in 40 iterations for the single object picture and in about 150-200 iterations for the two face object and the cube at each hierarchical stage.

6. CONCLUSIONS

The present paper has approached the object reconstruction and depth extraction from 3D-integral images as an inverse problem, which proves to be ill-conditioned. In order to cure the ill-conditioning of the problem, a hierarchical adaptive constrained 3D-reconstruction algorithm has been considered. It is based on the use of two sigmoid functions to determine a bounded constrained least squares solution which provides information about the number of objects in the scene, their shape and absolute and relative position. The scheme can be hierarchically reapplied to the relevant regions of the object space in order to produce higher resolution reconstructions. The simulation results have shown that the number of objects, absolute and relative positions and shape are recovered accurately. The hierarchical scheme produces high resolution reconstructions and is computationally efficient, as only the relevant regions of the scene are oversampled.

7. REFERENCES

[1] G. Lippmann, "La Photographie integrale", Comtes Rendus, Academie des Sciences, vol.146, pp. 446-451, 1908.

[2] T. Okoshi, "Three dimensional imaging techniques", Academic Press, London, 1976.

[3] N. Davies, M. McCormick, "Holoscopic imaging with true 3D-content in full natural colour", J.Phot.Science, vol. 40, pp. 46-49, 1992.

[4] N. Davies, M. McCormick, M. Brewin, "Design and analysis of an image transfer system using microlens arrays", Opt. Eng., vol. 33-11, pp. 3624-3633, 1994.

[5] S. Manolache, A. Aggoun, M. McCormick, N. Davies, "A mathematical model of a 3D-lenticular integral recording system", Proceedings of IEEE Vision, Modeling and Visualization, pp. 51-58, Erlangen, 1999.

[6] M. Bertero, P. Boccacci, "Introduction to inverse problems in imaging", Institute of Physics Publishing, Bristol and Philadelphia, 1998.

[7] S. Manolache, M. McCormick, S.-Y. Kung, "3D-object space reconstruction from planar recorded data of 3D-integral images", submitted for publication in Journal of VLSI Signal Processing Systems (Kluwer).

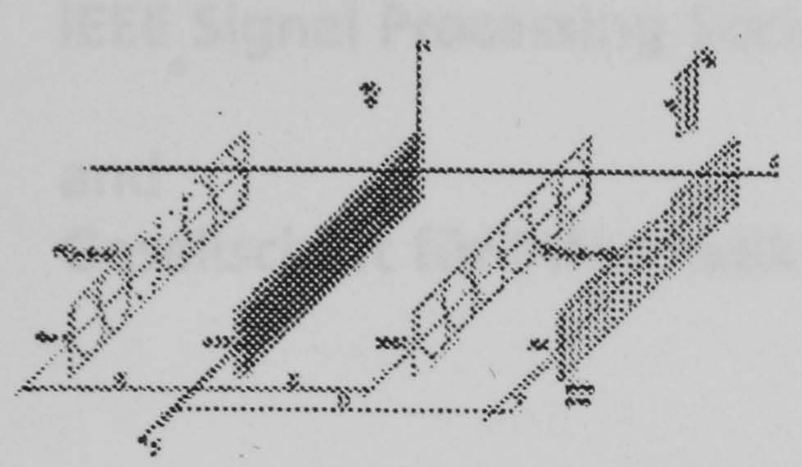


Fig. 1. 3D-lenticular integral imaging camera system

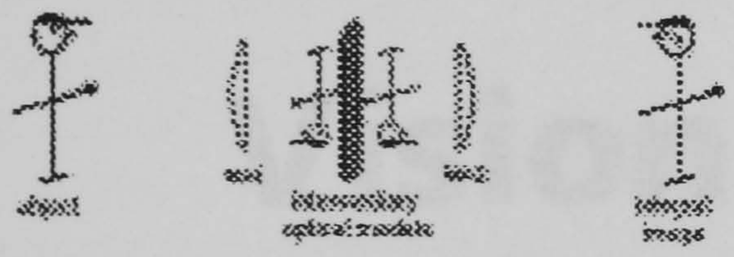


Fig. 2. Optical stages in the integral imaging

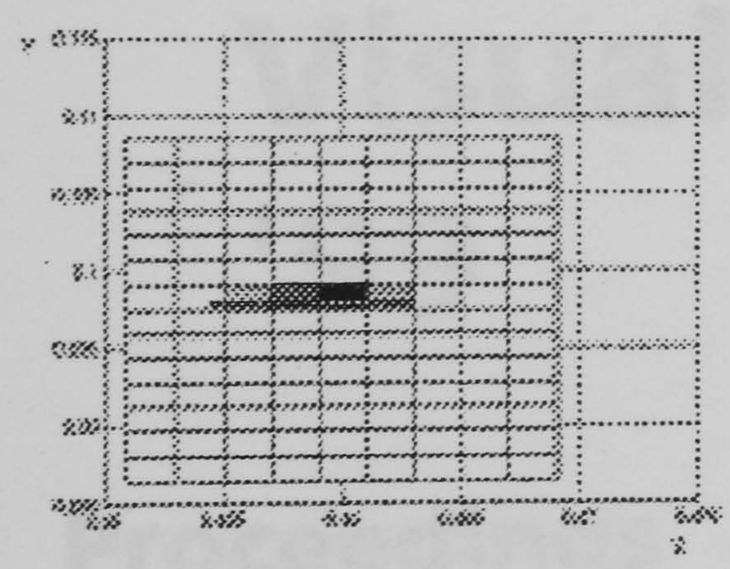


Fig. 3. Transversal section through the reconstruction of a thin face orthogonal to the recording screen. The original object is drawn in continuous line.

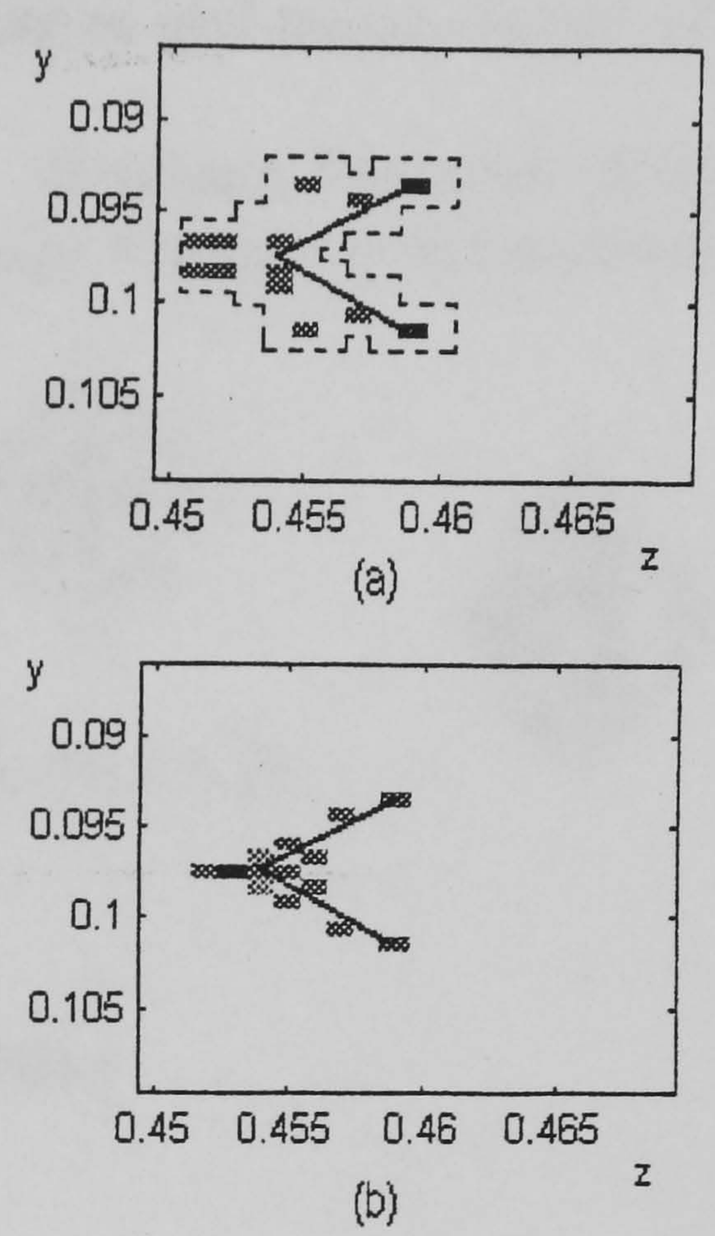


Fig. 4. Transversal sections through the hierarchical reconstructions of a cube. The original object is drawn in continuous line.

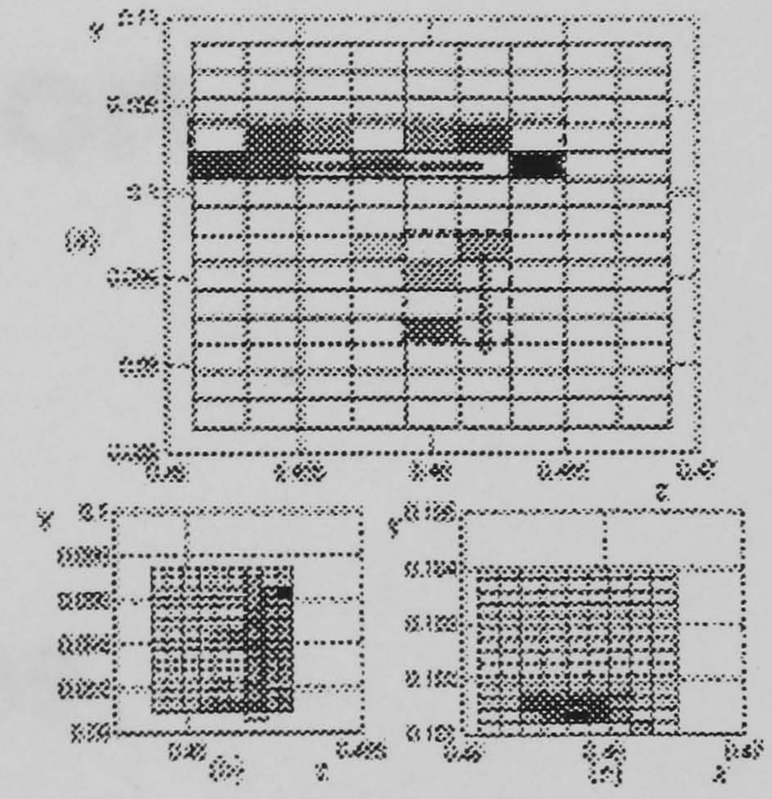
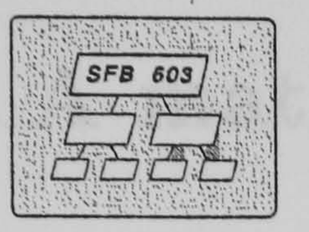


Fig. 5. Transversal sections through the hierarchical reconstructions of a pair of thin faces, one orthogonal and another parallel to the recording screen. The original objects are drawn in continuous line.



SFB 603 "Model-Based Analysis and Visualization of Complex Scenes and Sensor Data" and

Graduate Research Center "3-D Image Analysis and Synthesis"

in cooperation with
IEEE Signal Processing Society

and
Gesellschaft für Informatik GI (FG 1.0.4, FG 4.1.2)

B. Girod, H. Niemann, H.-P. Seidel (Eds.)

Vision Modeling and Visualization '99

Proceedings
November 17-19, 1999
Erlangen, Germany

infix

A mathematical model of a 3D-lenticular integral recording system

Silvia Manolache, Amar Aggoun, Malcolm McCormick, Neil Davies

De Montfort University, Department of Electronic and Electrical Engineering
The Gateway, Leicester LE1 9BH, UK
Email: silvia@dmu.ac.uk

Abstract

The paper presents a mathematical model of the 3D-lenticular integral recording camera system. It comprises a detailed analysis of the image formation, transmission, and recording phases, which leads to a geometrical model being developed. The coordinates of the intermediary, integrated and recorded images are established and a method by which the spatial position of a physical point can be derived from two recorded disparate micro-intensity distributions is presented. The number of recorded micro-intensity distributions representing a single object point and the dependence on spatial position is discussed and analysed. Point spread function for different segments of the camera and for the entire microlens system, using physical optics and mathematical analysis tools, is computed. In addition, the relationship between point spread dimensions and point depth is derived, and illustrative simulation results are presented.

1 Introduction

The development of 3D-imaging systems has been a constant pursuit of the scientific world in the XX-th century, involving a wide range of research areas. Integral photography was pioneered by Lippmann (1908), who used microlens arrays to create, record on photographic film, and replay integral three dimensional images. Since 1908, when it was first reported, the integral photographic technique has been improved as a result of theoretical studies, technical innovations of the optical systems, and progress in microlens manufacturing. It is now a viable alternative to other autostereoscopic systems.

The optics of an advanced form of integral imaging system – employing a two tier optical network – in which a true 3D optical reconstruction of a scene is transferred to the capture device, has been described in detail by Davies and McCormick ([1], [2], [3]). The optical arrangement, shown in figure 1, comprises two macrolens arrays (MA1 and MA2) placed equidistantly behind and in front of an autocollimating transmission screen (ATS). The ATS is made up of two microlens arrays separated by their joint focal distance. The recording plane, a photographic plate whose position coincides with the focal plane of a microlens array (RA), lies within the optical model.

Within this paper, each object point is represented by a number of information distributions: an *image* is defined to be a point by point reconstruction of the object space; an *optical model* is considered to be formed by intersecting modulated ray bundles; the final recorded data is a *sampled set of disparate micro-intensity distributions.*

The object is imaged by the input macrolenses which transmit compressed transposed images that occur at or near the central double microlens screen (ATS). The screen inverts the optical sense of each interme-

diary image (see figure 2), and, simultaneously, presents these spatially reversed 3D-optical models to the corresponding output macrolenses.The output macrolenses retranspose the optical models to the correct spatial location. The final integrated image, formed by superposed optical models projected by the second macrolens array, is a true 3D optical 1:1 reconstruction of the object. This can be recorded as a sampled data set possessing either horizontal parallax (if semicylindrical microlenses have been used in the recording array), or continous parallax (in the case of the spherical or square based microlens arrangement).

The integral image recorded in the focal plane of a recording microlens array as a planar sampled data set contains all the 3D information related to the object space. Each microlens of the recording array samples a fractional part of the scene, many microlenses recording directional information of the scene from different viewing angles. Therefore, parallax information about any particular point is spread over the recording plane. Redisplay of the full spatial model as a real 3D image can be effected by overlaying the sampled data set by an integral decoding element.

The paper presents a mathematical model of the integral imaging camera system for the case where the ATS comprises square based lenslets and the recording array is made up of identical semicylindrical microlenses. This system is known as a *3D-lenticular integral imaging system* ([4]) and produces 3D images containing horizontal parallax.

2 Geometrical aspects of integral image formation and recording

The geometrical aspects of integral image formation and recording are described in full detail. The coordinates of the intermediary and final optical model, and of the recorded data are computed using geometrical optics laws.

All computations and formulae are related to the Cartesian coordinate system $Oxyz$ depicted in figure 1. The z-axis denotes the depth direction, while the x, y-axes describe the lateral positions. The macro- and microlenses are indexed in the sense indicated by the positive axis Oy.

Any imaged point $P(x_p, y_p, z_p)$ has a number of states prior to the recording stage. The macrolenses of the first array present a first set of images of P near the autocollimating transmission screen. These images are located at the points P_k^1 having coordinates:

$$x_k^1 = \frac{x_P f_1}{f_1 + S + z_P}$$

$$y_k^1 = \frac{b_k z_P + b_k S + y_P f_1}{f_1 + S + z_P}$$

$$z_k^1 = \frac{z_P f_1 - z_P S - S^2}{f_1 + S + z_P},$$

for any macrolens k in the first array, where: f_1 is the focal length of the macrolenses, b_k denotes the y-coordinate of macrolens k 's centre, and S is the distance between a macrolens array and the central plane Oxy (see figure 1 and figure 3(a)).

These intermediary images become object points for the ATS. The light emerging from them, refracted and diffracted through the first microlens array, produces an intensity distribution in the common focal plane of the two microlens arrays; this intensity distribution is transmitted by the second microlens array through a symmetric optical process. As a result, the ATS operates by refracting the transmitted ray at an equal angle to the angle of incidence of the impinging ray. In this way, every "object" point P_k^1 in front of the screen has a quasi-equidistant voxel representation centred at the point $P_k^2(x_k^2, y_k^2, z_k^2)$ behind the screen, namely:

$$x_k^2 = \frac{x_P f_1}{f_1 + S + z_P}$$

$$y_k^2 = \frac{b_k z_P + b_k S + y_P f_1}{f_1 + S + z_P}$$

$$z_k^2 = \frac{-z_P f_1 + z_P S + S^2}{f_1 + S + z_P}$$

Each intermediary voxel P_k^2 is collected by the corresponding output macrolens k and imaged at the location $P'(x_P, y_P, z_P)$ in space, which is the equal conjugate image location of point P with respect to the Oxy plane. It does not depend on the macrolens k, and therefore the result is an integration of directional information about P from all the macro- and microlenses that have imaged it.

The intensity distributions related to the integral image P' of P are recorded on a photographic plate that lies behind the recording microlens array (containing microlenses of pitch ϕ_r, length ψ_r, and focal length F, situated at a distance D from the central Oxy plane - figure 1). P' has recorded intensity distributions in a number of microlenses, the number being determined by the axial position of P'.

If microlens k "sees" the point P', then the recorded data behind microlens k is centred at the point P'_k whose coordinates are:

$$x'_k = x_P$$

$$y'_k = \frac{(c_1 + \phi_r k)(D + F - |z_P|) - y_P F}{D - |z_P|}$$

$$z'_k = D + F\,,$$

where c_1 is the ordinate of the first microlens centre.

The disparity between two recorded intensity distributions of the point P corresponding to microlenses k and j is:

$$d_{jk} = \frac{(D + F - |z_P|)|j - k|\phi_r}{D - |z_P|}\,.$$

This expression allows the recovery of the physical point's position when two of its intensity distributions, centred at $P'_j(x'_j, y'_j, D + F)$ and $P'_k(x'_k, y'_k, D + F)$, are known (and, therefore, the disparity d_{jk} between them is known). Namely:

$$x_P = x'_j = x'_k$$

$$y_P = \frac{(c_1 + \phi_r j)d_{jk} - y'_j|j - k|\phi_r}{D - |z_P|}$$

$$z_P = \frac{(D + F)|j - k|\phi_r - d_{jk}D}{d_{jk} - |j - k|\phi_r}\,.$$

Another important problem concerns the number of recorded intensity distributions of the optical reconstruction P' of a physical point P (figure 3(b)). Computations prove that this number depends on both lateral position and depth of the point. The maximum recorded image number for a given depth is attained for points having a central position in the imaged scene:

$$N_{max} = \left[ap\frac{D - |z_P|}{\phi_r(|z_P| - S)}\right]\,.$$

This number depends essentially on the camera aperture ap. Marginal points have obviously less recorded images due to the finite dimensions of the recording screen. A Matlab application plots the step functions that describe the dependence between the number of recorded intensity distributions of a point and depth for different lateral positions of the point. Figure 4 depicts the number of recorded intensity distributions for marginal points (a), respectively central points (b), with depth ranging between 0.4 m and 0.58 m when the recording screen is at a distance D=0.58 m from the central plane of a camera having an aperture ap=0.3 m. As can be seen, points in the proximity of the recording plane are not recorded. This singularity is explained in terms of physical optics in the next section.

3 Physical aspects of the optical process

3.1 Optical imaging model

Due to the very fine pitch of a microlens, diffraction plays a major role during the optical processes which lead to the integral image. Therefore, a model of image formation based on the ideas of Rayleigh and Duffieux, which visualises an image as the combination of the diffraction patterns produced by every individual point in the object, has been considered ([5]). In this model, each point in an object being imaged is itself considered as a point source of light. The wavefronts from

it are limited in their entry into the imaging system by the finite aperture of the imaging lens, and the diffraction pattern of that aperture is formed in the conjugate plane to the plane containing the point source. Each point is therefore imaged not as a point, but as the diffraction pattern of the aperture of the imaging lens or more complicated patterns if aberrations are present. The centre of the diffraction pattern of a point is, as expected, its conjugate according to the geometrical optics. The image of a continuous object distribution is expressed as a convolution of that distribution with *the point spread function* (PSF), or impulse response of the system.

3.2 Point spread function in the 3D-lenticular integral recording system

As a result of the optical process, each point will give rise to an intensity distribution in the image: this is *the point spread function.* It will be the same for all object points in a sufficiently small neighbourhood. The point spread function is defined ([6]) as the square of the modulus of the complex amplitude of the light in a reference plane spanned by an XY Cartesian frame:

$$psf(X,Y) = |A(X,Y)|^2 .$$

The calculation of the point spread function of the lenticular integral imaging system has to take into account the three microlens arrays through which light passes prior to the integral reconstruction and recording of the imaged object. Using the Fresnel - Kirchhoff equation, Burckhardt ([7], [8]) derived the formula of the point spread function for a microlens screen along the lateral axis Ox. Adapting his result to the lenticular integral camera system, which uses square based lenslets in the ATS, the point spread function of the first microlens array in the focal plane is:

$$psf_1(X,Y) = \exp\left(-\frac{X^2+Y^2}{u^2}\right) \qquad (1)$$

where the coefficient:

$$u^2 = 0.106\frac{\phi^2 b^2 f^2}{a^4} + 0.245\frac{\lambda^2 f^2}{\phi^2},$$

depends on camera parameters and the position of a point in space. In the expression, ϕ is the pitch, f is the focal distance of a microlens, b is the thickness of a relevant portion of scene which contains point P, and a is the distance between an intermediary image and a microlens array, given by:

$$a = \frac{-z_P f_1 + z_P S + S^2}{f_1 + z_P + S} - f .$$

The thickness parameter b can be chosen in accordance with the depth resolution of the image. In all simulations presented in this paper, b has been assumed to be equal to 1 mm. Analysing u^2, it can be noticed that the first term of the sum expresses the defocus influence, whilst the second one is due to diffraction ([4]). Due to the symmetry of the optical process that occurs at the central double microlens array, the point spread function of the second microlens array is also given by (1). Therefore, the spread function of the spatial inversion process (associated to the central screen) is given by the convolution of (1) with itself:

$$psf_{scr} = psf_1 \star psf_1 .$$

A straightforward computation yields for psf_{scr} the next Gaussian distribution:

$$psf_{scr}(X,Y) = v^2 \exp\left(-\frac{X^2+Y^2}{v^2}\right), \qquad (2)$$

where:

$$v^2 = 2u^2 .$$

The point spread function associated with the recording semicylindrical microlens screen in the focal plane (which is the plane of the photographic plate) has an expression analogous to (1), but containing different parameters:

$$psf_{red}(X,Y) = \exp\left(-\frac{X^2}{w_1^2}\right)\exp\left(-\frac{Y^2}{w_2^2}\right), \qquad (3)$$

where:

$$w_1^2 = 0.106\frac{\psi_r^2 b^2 F^2}{a_r^4} + 0.245\frac{\lambda^2 F^2}{\psi_r^2}\,,$$

$$w_2^2 = 0.106\frac{\phi_r^2 b^2 F^2}{a_r^4} + 0.245\frac{\lambda^2 F^2}{\phi_r^2}\,,$$

In the above formulae, ψ_r, ϕ_r, and F are the length, width, and the focal distance of a recording microlens, whilst a_r is the distance between the reconstructed model P' and the recording microlens array ($a_r = D - |z_P|$), and λ is the wavelength.

The spread function of the entire optical process combines the spread effect of both the central double microlens screen and the recording array. Therefore, it is the convolution of psf_{scr} and psf_{rec}:

$$psft = psf_{scr} \star psf_{rec}\,.$$

Computations yield:

$$psft(X,Y) = K\exp\left(-\frac{X^2}{a_1^2}\right)\exp\left(-\frac{Y^2}{a_2^2}\right), \quad (4)$$

where

$$K = \frac{v^4 w_1^2 w_2^2}{\sqrt{(v^2+w_1^2)(v^2+w_2^2)}}\,,$$

$$a_1^2 = v^2 + w_1^2\,, \qquad a_2^2 = v^2 + w_2^2\,.$$

Formula (4) expresses the spread function in local coordinates associated with a microlens. Using the global Cartesian system depicted in figure 1, the point spread behind a specific microlens k of the recording array, imaging a point P can be written as:

$$psft_k(x,y) = K\exp\left(-\frac{(x-x'_k)^2}{a_1^2}\right)\exp\left(-\frac{(y-y'_k)^2}{a_2^2}\right),$$

where x'_k, and y'_k are the coordinates of the recorded intensity distribution P'_k behind microlens k, given in the previous section.

The Fresnel - Kirchhoff formula, which is not a rigorous deduction from the wave equation, but an approximation, can be improved by considering a correction, the inclination factor $Q = \frac{1}{2}(\cos\theta_n + \cos\theta'_n)$, used under the integral in the Fresnel - Kirchhoff equation [9]. The angles θ_n and θ'_n, shown in figure 6(a), are the angles that the incident and diffracted ray would make with the normal unit vector $\bar{n}$ of the diffracting aperture (the microlens, in the present case). The inclination factor Q is a slowly varying function of θ_n and θ'_n and, as the microlens aperture is small, it can be treated as a constant over the surface of one microlens. Considering the case in figure 6(b), $\theta_n = \theta'_n = \theta_k$, and thus $Q = \cos\theta_k$, where θ_k is the angle subtended to the normal by the ray $P'P'_k$. Using this inclination factor in the previous computations, the point spread function of the whole process becomes:

$$psf_{total}(x,y) = \sum_k \cos^2\theta_k\, psft_k(x,y)\,,$$

where the summation is done with respect to the indices k of the microlenses involved in recording point P.

The shape of $psft$, which is a two variable Gaussian function, suggests that the energy forming the recorded intensity distribution is concentrated on a rectangular spot whose dimensions depend on point depth. Matlab applications have been used to plot the dependence on depth for both the width and the length of a rectangle that concentrates more than 98% of the intensity. Figure 5 presents results corresponding to a recording screen position D =0.58 m. The graphs prove that the dimensions attain a stability level for reconstructed points situated at more than 3 cm away from the recording screen, and that the spread is larger when the recorded points are nearer the recording screen. For the points reconstructed in the immediate vicinity of the recording plane (at a distance smaller than 1.5 mm) , the spread is infinite, therefore these points do not contribute stricto sensu to the recorded spatial information. This result confirms that the number of contributing intensity distributions for these points is 0, as proved on a geometrical basis in the previous section.

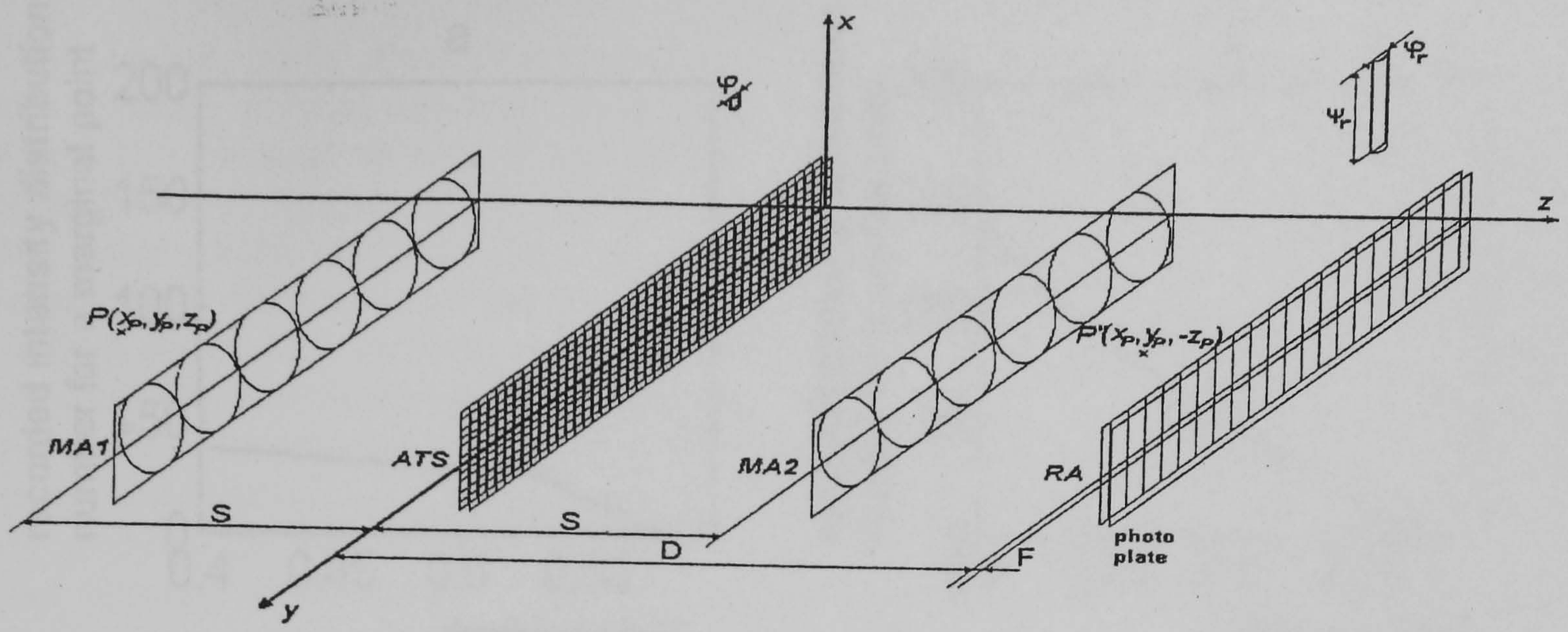


Figure 1: 3D-lenticular integral imaging camera system

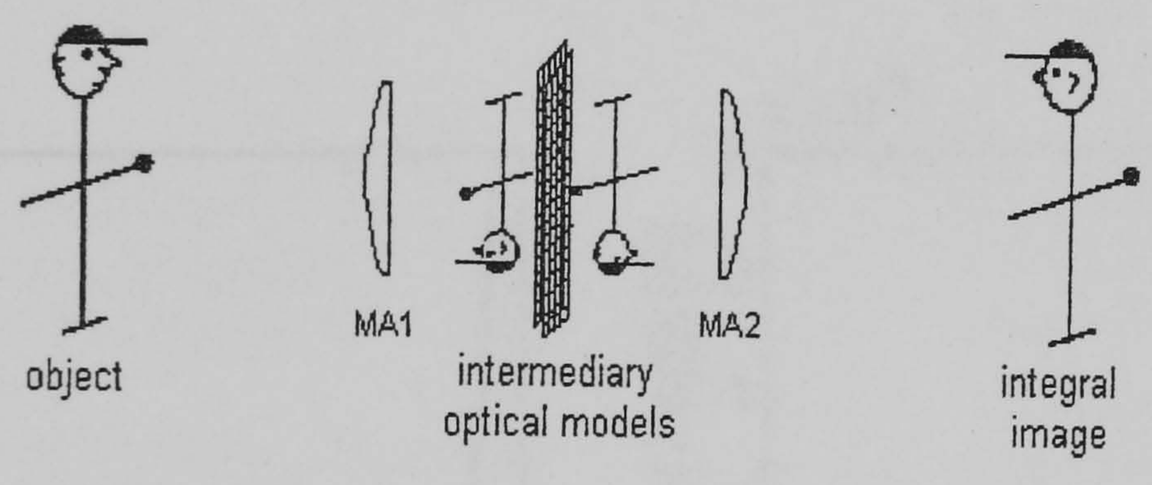


Figure 2: Optical stages in the integral imaging

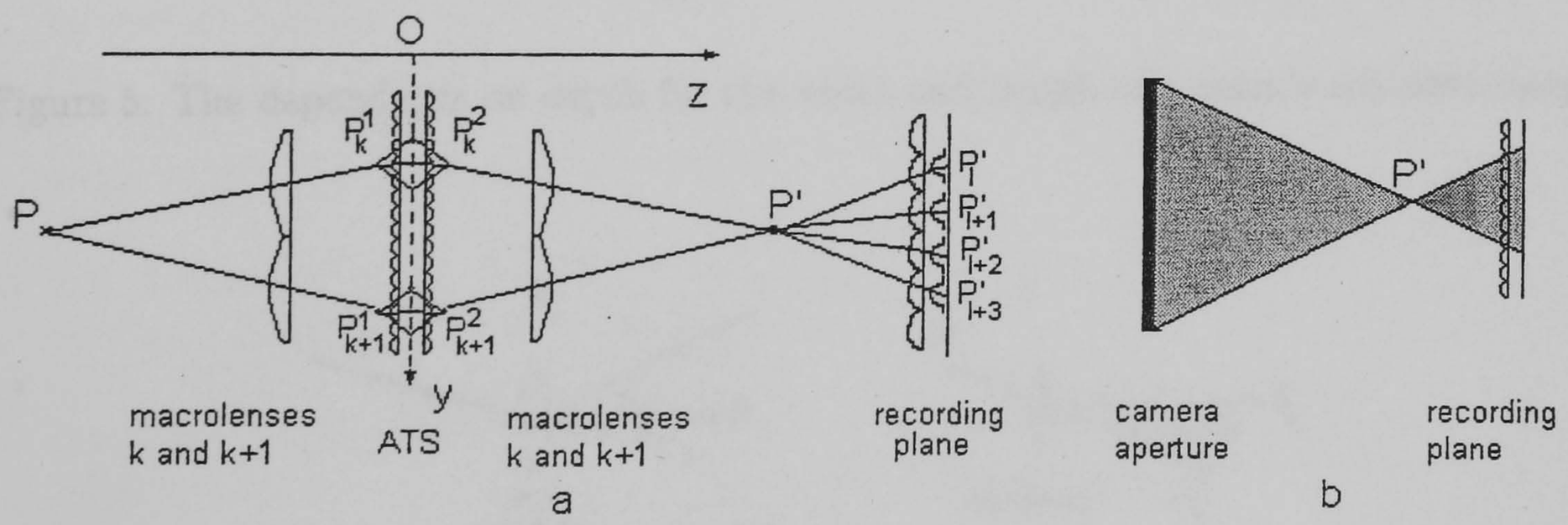


Figure 3: Transverse sections of the arrangement in figure 1 with a plane $x = x_P$: (a)Schematic representation of the intermediary, integral and recorded optical models of a physical point P; (b) The microlenses of the recording array that record a certain point P'

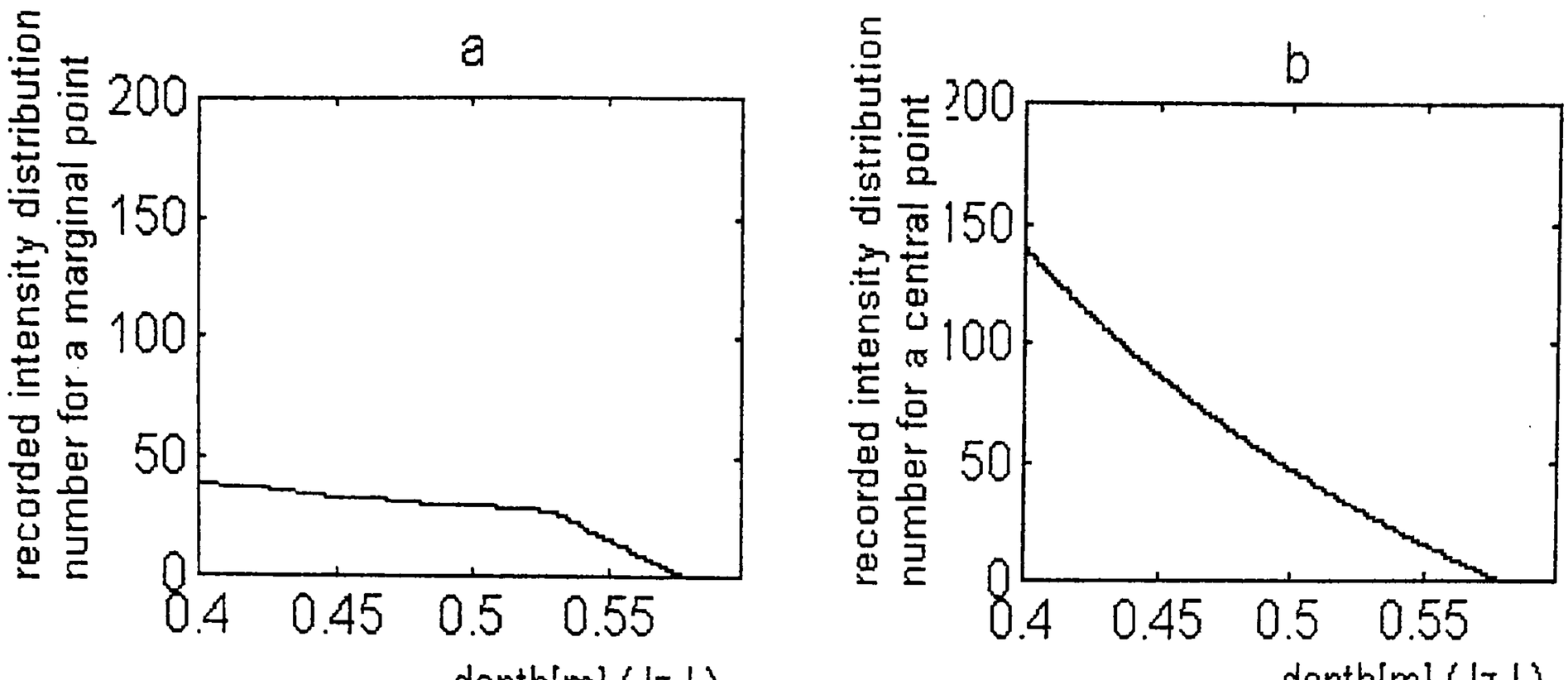


Figure 4: The dependence on depth of the number of recorded intensity distributions for a marginal (a), respectively a central point (b).

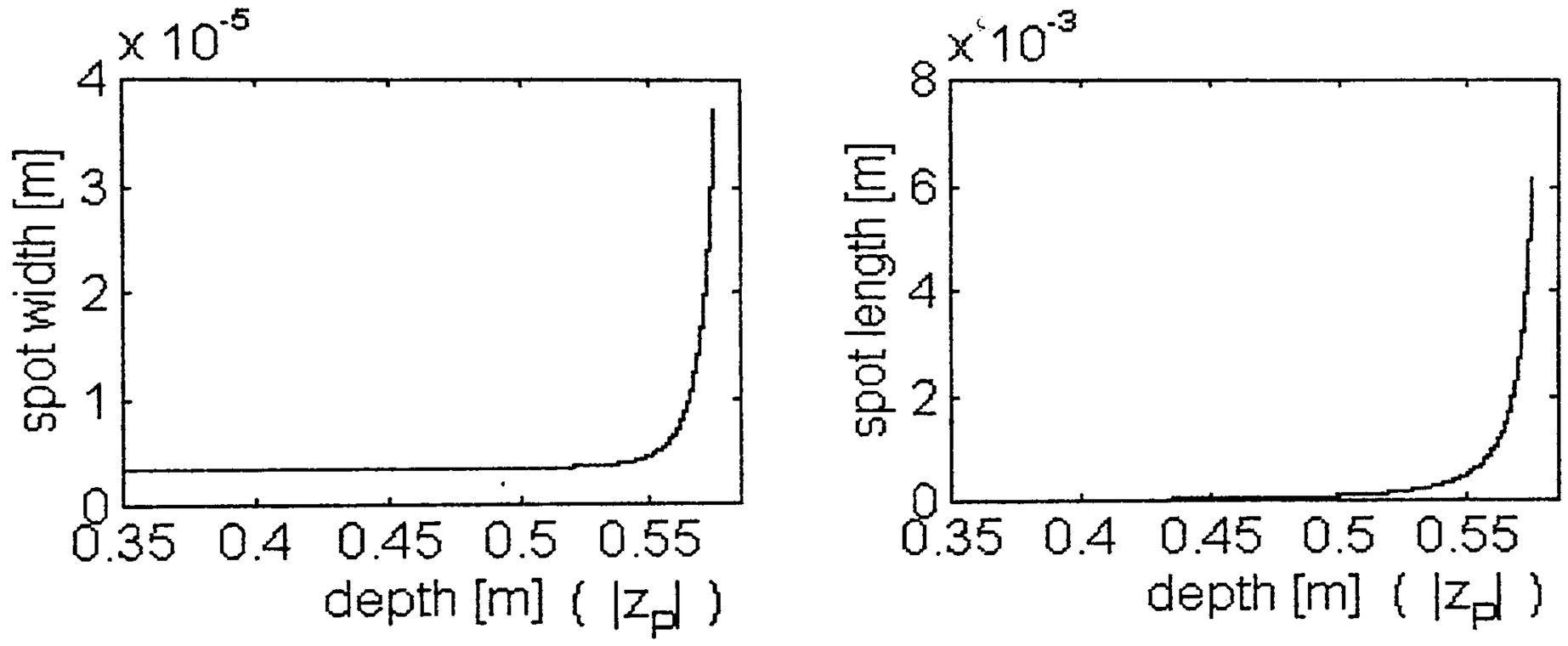


Figure 5: The dependence on depth for the width and length of a point's recorded image.

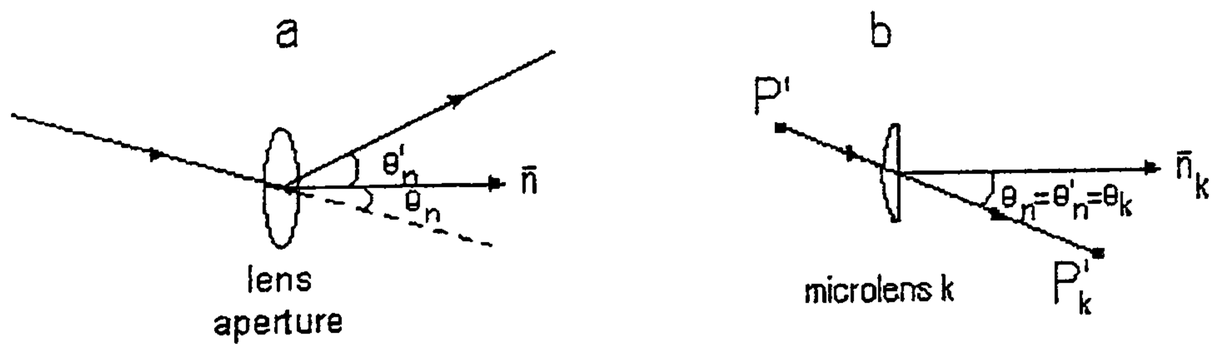


Figure 6: Incidence and diffraction angle at the diffraction through a lens aperture: a. general case; b. particular case: diffraction through a microlens aperture

4 Conclusions

A model of the 3D-lenticular integral recording camera system, involving both geometrical and physical aspects of the optical process, has been described. The coordinates of the intermediary, integrated, and recorded data have been computed and used to retrieve an object point position from two recorded micro-intensity distributions. The problem of computing the number of recorded micro-intensity distributions and the dependence on point position has been treated. The point spread function of the entire process and its dependence on point depth has been analysed, and simulation results are presented. These illustrate the dependence of recorded intensity distributions and object footprint on spatial position.

References

[1] N. Davies, M. McCormick, "Holoscopic imaging with true 3D-content in full natural colour",J.Phot.Science, vol.40, pp.46-49, 1992.

[2] N. Davies, M. McCormick, "Holoscopic imaging with true 3D-content in full natural colour", J.Phot.Science, vol.40, pp.46-49, 1992.

[3] N. Davies, M. McCormick, M. Brewin, "Design and analysis of an image transfer system using microlens arrays", Opt. Eng., vol.33, no.11, pp.3624-3633, 1994.

[4] T. Okoshi, "Three dimensional imaging techniques", Academic Press, London, 1976.

[5] E.G. Steward, "Fourier optics - an introduction", Ellis Horwood ltd., 1987.

[6] K.R. Barnes, "The Optical Transfer Function", Adam Hilger ltd, London, 1971.

[7] C. Burckhardt, "Optimum parameters and resolution limitation of integral photography", J. Opt. Soc. Am., vol.58, no.1, pp.71-76, 1968.

[8] C. Burckhardt, R.J. Collier, E.T. Doherty, "Formation and Inversion of pseudoscopic images", Appl. Opt., vol.7, no.4, pp.627-631, 1968.

[9] M.V. Klein, T.E. Furtak, "Optics", John Wiley&Sons, New York, 1986.

NUMÉRISATION 3D

SCANNING 2001

Le Rendez-vous Européen Annuel de la Numérisation
Sixième Edition

ACTES DES CONFERENCES
NUMERISATION 3D

Mercredi 4 et jeudi 5 avril 2001

Union Internationale des Chemins de Fer (UIC)
Paris 15ème

Avec le Parrainage de :
Contrôle Industriel - Pixel – Photoniques – Wholers associates – Produits équipements industriels (PEI)

Regularisation Methods for Object Space Reconstruction from 3D-Integral Images

Méthodes de régularisation pour la reconstruction de l'espace-objet d'une image intégrale tridimensionelle

Silvia Manolache[1], S.Y. Kung[2], Malcolm McCormick[1], Amar Aggoun[1]

[1]De Montfort University, Department of Engineering and Technology, The Gateway, Leicester LE1 9BH, UK, Tel: +44-(0)116-2551551 x8082, Fax: +44-(0)116-2577692, E-mail: silvia@dmu.ac.uk

[2]Princeton University, Department of Electrical Engineering, Princeton, NJ 0841

Abstract

The paper presents two methods for object space reconstruction from planar recorded data of 3D-integral images. The integral imaging system is described in detail and the associated point spread function is given. The depth estimation from 3D-integral pictures is formulated as an inverse problem: given the image, i.e. the recorded 'effect' of the object space , find the 'cause' which had produced it , i.e. recover the intensity map and, consequently, the composition and the depth of the object scene. In order to cure the ill-posedness of the problem, approximate solutions satisfying additional constraints coming from the physics of the problem are searched using so called 'regularisation methods'. The present work comprises two regularisation schemes for obtaining constrained least-squares solutions of the object space reconstruction problem. The first algorithm is based on the projected Landweber method. The second method is based on the iterative use of sigmoid functions to obtain the object space reconstruction as a positive least-squares solution of the imaging equation. A hierarchical multiresolution strategy of applying the regularisation methods, providing higher resolution reconstructions, is also proposed. Finally, illustrative simulation results are given.

Resume

L'article présente deux méthodes de reconstruction de l'espace-objet à partir d'une image integrale à 3D, enregistrée sous la forme de données planaires (bidimensionelles). Le système d'imagerie intégrale est décrit en detail et la fonction d'étalement du point (PSF) associée est donnée. L'estimation de la profondeur d'une image intégrale est formulée comme un problème invers: en donnant l'image, qui est l'*effet* enregistré de l'espace-objet, on cherche la *cause* qui l'avait produit, i.e. l'intensité dans l'espace-objet et, par conséquence, la composition et la profondeur de la scène. Pour traîter le mal conditionnement du problème, on cherche des solutions approximatives qui satisfaient des conditions supplémentaires provenant de la realité physique, à l'aide des methodes de régularisation. Le premier algorithm est basé sur la méthode de Landweber projetée. Le second est un schéma iteratif qui emploît des fonctions sigmoïdes pour obtenir des reconstructions de l'espace-objet sous la forme de solutions positives du problème des moindres carrés associé à l'équation de l'imagerie. Une stratégie hiérarchique multirésolutionelle d'appliquation des méthodes de régularisation, qui permet une reconstruction à meilleure résolution, est proposée. Finalement, on présente des résultats de simulation pour illustrer les deux méthodes.

1. Introduction

Integral photography was introduced in 1908 by Lippmann ([1]), who used microlens arrays to create, record on photographic film, and replay integral three dimensional images. Integral imaging resembles to holography, but it uses natural light and reproduces true colour optical models. Hence, it offers a viable alternative to other autostereoscopic systems ([2]).

The optics of an advanced form of integral imaging system -- employing a two tier optical network -- in which a true 3D optical reconstruction of a scene is transferred to the capture device, has been described in detail by Davies and McCormick ([3], [4]).

Extracting depth information from 3D-integral images has various applications, which range from remote inspection in robotic vision and medical imaging to combining real and computer generated 3D-pictures in a virtual studio. The aim of the present paper is to describe a method of reconstructing the composition of a 3D-object space from 2D-recorded data encoding the scene. The object space reconstruction from 3D-integral pictures is formulated as an inverse problem: given the image, i.e. the recorded 'effect' of the object space, find the 'cause' which had produced it, i.e. recover the intensity contours and, consequently, the composition of the object scene. The direct problem - image formation and recording - has been studied in a previous paper, and the equation of the point spread function for the lenticular integral imaging system has been derived ([5]).
Inverse problems in imaging are ill-posed and their discrete correspondents are ill-conditioned due to the inherent loss of information associated with the direct process ([6]). In order to cure the ill-posedness of the problem, approximate solutions satisfying additional constraints coming from the physics of the problem are searched.

The present work comprises two regularisation schemes for obtaining a constrained least-squares solution of the scene reconstruction problem, which are hierarchically applied in order to obtain high resolution object space reconstructions.

2. 3D-integral imaging system and associated point spread function

The optical arrangement of the 3D-integral imaging system, shown in Figure 1, comprises two macrolens arrays (MA1 and MA2) placed equidistantly behind and in front of an autocollimating transmission screen (ATS). The ATS is made up of two microlens arrays separated by their joint focal distance. The recording plane, a photographic plate whose position coincides with the focal plane of a microlens array (RA), lies within the optical model.

The paper considers a camera configuration where the ATS contains square based hemispherical lenslets and the recording array is made up of identical semicylindrical microlenses. This system is known as a *3D-lenticular integral imaging system* ([3], [5]) and produces 3D-images containing horizontal parallax.

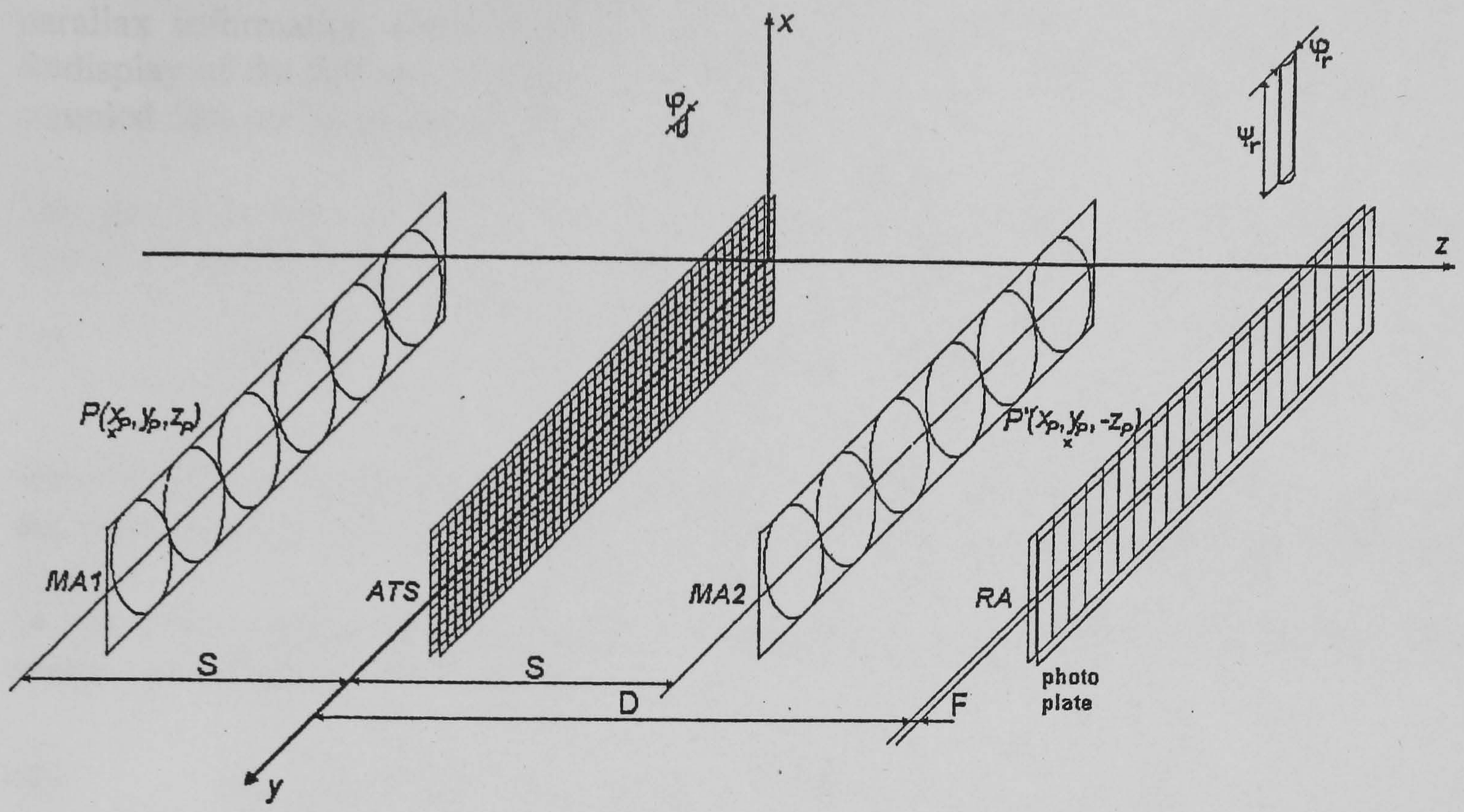


Figure 1. Integral imaging camera system schematic.

The object is imaged by the input macrolenses which transmit compressed transposed images that occur at or near the central double microlens screen (ATS). The screen inverts the optical sense of each intermediary image (see Figure 2), and, simultaneously, presents these spatially reversed 3D-optical models to the corresponding output macrolenses. The output macrolenses retranspose the optical models to the correct spatial location. The final integrated image, formed by superposed optical models projected by the second macrolens array, is a true 3D optical 1:1 reconstruction of the object. It is recorded as a *2D-sampled data set* possessing either horizontal parallax (if semicylindrical microlenses are used in the recording array), or continuous parallax (in the case of the spherical or square based microlens arrangement).

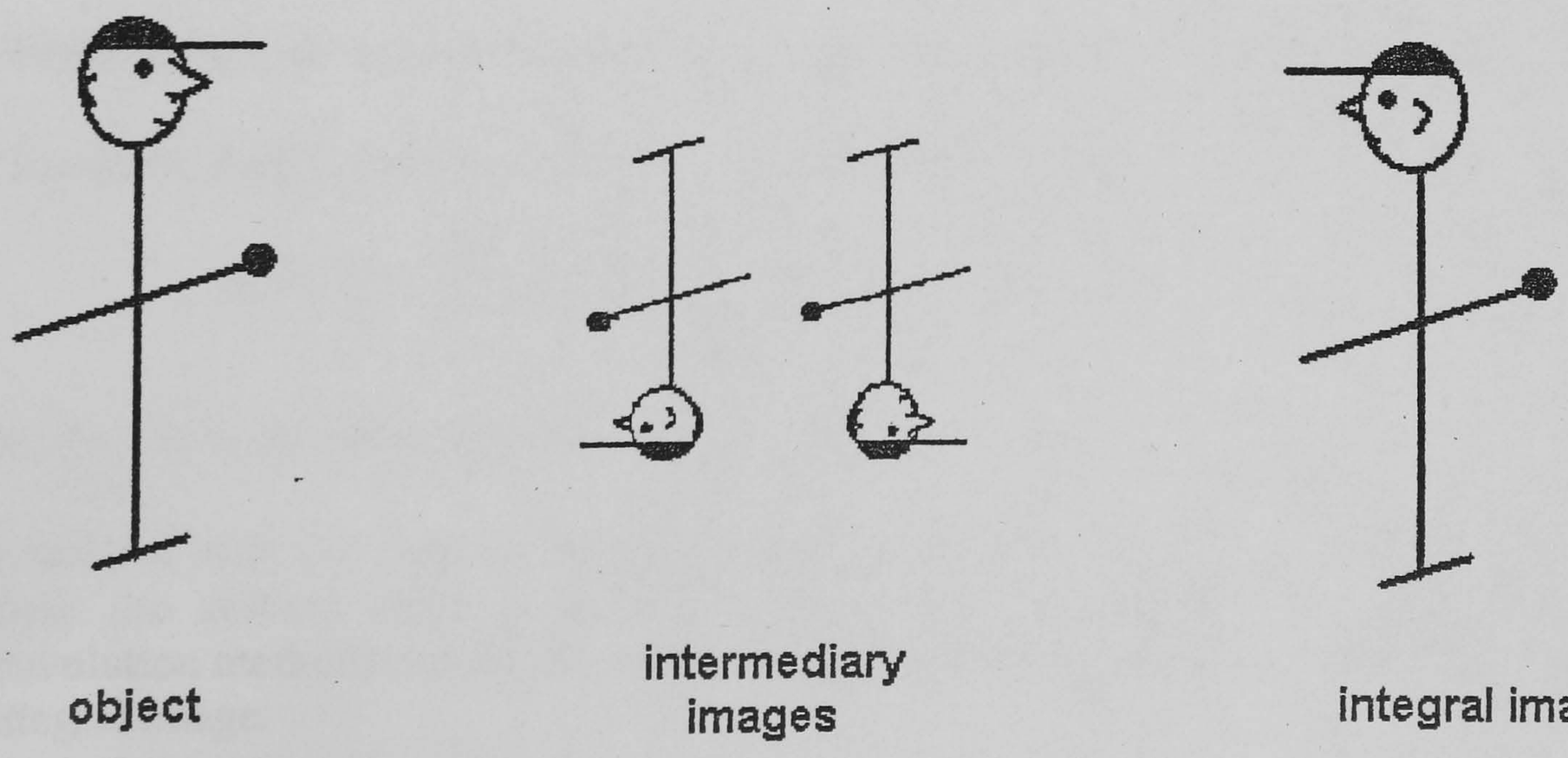


Figure 2. Operation of the integral camera system.

The integral image recorded in the focal plane of a recording microlens array as a *planar* sampled data set contains all the 3D information related to the object space. Each

microlens of the recording array samples a fractional part of the scene, many microlenses recording directional information of the scene from different viewing angles. Therefore, parallax information about any particular point is spread over the recording plane. Redisplay of the full spatial model as a real 3D image can be effected by overlaying the sampled data set by an integral decoding element.

The spread function of the entire optical process combines the spread effect of both the central double microlens screen and the recording array ([5]):

$$K(X,Y,x,y,z) = \sum_{l(x,y,z)} K_{l(x,y,z)}(X,Y,x,y,z) \tag{1}$$

where *(X,Y)* are the coordinates in the image space, *(x,y,z)* represent the object space, and the summation is done with respect to the index of the microlenses which 'see' the point *(x,y,z)*.
In the above equation, $K_{l(x,y,z)}$ is the point spread function component behind one microlens of the recording screen:

$$K_{l(x,y,z)}(X,Y) = \alpha(z)\ \exp\left(-\frac{(X-X_l)^2}{a_1^2(z)}\right)\ \exp\left(-\frac{(Y-Y_l)^2}{a_2^2(z)}\right) \tag{2}$$

The point (X_l, Y_l) represents the point of maximum intensity in the microimage formed behind microlens l, and a_1, a_2, α depend explicitly on lenslet parameters and point depth z ([5]).

Equation (1) indicates that the spread function is space variant and that it holds depth information in all the factors. This fact will be used in recovering three dimensional data about the object space.
Due to the small values of the parameters involved in the point spread function formulae, only the significant factors should be considered in a computational approach of the depth extraction problem. When using lenticular arrays (i.e. semicylindrical microlenses), the spread on the x-dimension decreases rapidly to 0, so the x-factors in the point spread expression can be dropped.

3. Formulation of the object space reconstruction as an inverse problem

The image formation equation for the integral imaging system is:

$$g(X,Y) = \int I(x,y,z) \sum_{l(x,y,z)} K_{l(x,y,z)}\ dx\ dy\ dz \tag{3}$$

where $I(x,y,z)$ is the intensity at the point *(x,y,z)*.

It is noticed that the integral image formation process is not a convolution process because the critical shift invariance property does not hold. Therefore classical deconvolution methods cannot be used to tackle the problem of scene reconstruction from an integral image.

However, the point spread function provides a linear image formation operator A such that equation (\ref{imform1}) be written as:

(4) $$g = A\ I\ .$$

The reconstruction of the intensity distribution in the object is the inverse process of the image formation given by equation (4).
The discrete variant of the problem has been treated considering various samplings of the object space. Numerical experiments have shown that the matrix associated with the point spread function is very ill-conditioned. Alternative methods of reconstructing the object space are thus necessary and one of them is presented in the next section.

4. Regularisation methods for object space reconstruction

In order to overcome the ill-conditioning of this inverse problem, additional constraints coming from a priori knowledge about the object space have to be imposed upon the sought solution $\hat{I}$. More precisely, the solution $\hat{I}$ has to be such that $A\ \hat{I}$ is as close as possible to the image g, i.e.:

(5) $$\left\| A\ \hat{I} - g \right\| = \min .$$

Also, a condition of positivity and bounding is necessary, as there are neither negative nor infinite intensity objects:

(6) $$\hat{I} \in [0,T]^n .$$

a) Projected Landweber (PL) method

A solution satisfying conditions (5) and (6) can be regarded as a least-squares solution belonging to the closed and convex set C of the functions bounded by 0 and T:

(7) $$C = \left\{ \hat{I} \,\middle|\, 0 \le \hat{I} \le T \right\}$$

A method to obtain such a solution has been outlined by Landweber ([7]) and further studied by Eicke ([8]). Let us denote by $\mathcal{J} = \overline{A(C)}$ the closure of the set $A(C)$ and by $P_{\mathcal{J}}$ the convex projection onto $\mathcal{J}$. Then the constrained problem (5)-(6) has a solution if and only if $P_{\mathcal{J}}\ g \in A(C)$. In this case, the set of the constrained least-squares solutions is closed and convex. Eicke ([8]) proved that the set of the least-squares solutions coincides with the set of the fixed points of the non-linear operator $P_C G$, given by:

(8) $$P_C G(\hat{I}) = P_C(\hat{I} + \tau(A^* g - A^* A\hat{I})).$$

The projection operator P_C is non-linear and non-expansive. Therefore, for τ satisfying $0 < \tau < 2/\|A\|^2$, the compound operator $P_C G$ is also non-expansiv.
The projected Landweber (PL) method *is the iterative derivation of approximations of the fixed points of the operator* $P_C G$*:*

(9) $\hat{I}_{k+1} = P_C G(\hat{I}_k) = P_C(\hat{I}_k + \tau(A^* g - A^* A\hat{I}_k))$

starting with an arbitrary initial guess $\hat{I}_0$. From the general fixed point theorems for non-expansive operators, it yields that the sequence $\{\hat{I}_n\}_{n\geq 1}$ converges weakly to a fixed point of $P_C G$, which is a solution of the constrained problem (5)-(6).

b) Adaptive constrained 3D-reconstruction (AC3DR) algorithm

Another regularisation method can be derived if $\hat{I}$ satisfying (5)-(6) is written as the thresholded value of arbitrary data u: $\hat{I} = f(u)$, where f is a thresholding function. The projection onto the interval *[0,T]* is not differentiable, thus a sigmoid approximation of it, $f_\sigma : R^n \to [0,T]^n$, has been preferred:

(10) $$f_\sigma(u_i) = \frac{T}{1 + \exp\left(-\left(\frac{u_i}{\sigma} - k\right)\right)}, \quad i = \overline{1,n}.$$

where n is the number of components of the vector $\hat{I}$, i.e. the number of sampling points chosen in the object space. σ is an arbitrary number which controls the slope of the thresholding curve and k is a translation term ([9]).

To sum up, the sought vector $\hat{I}$ has the form $\hat{I} = f_\sigma(u)$ and minimises the discrepancy functional $\varepsilon = \|A\ I - g\|$.

The gradient of ε with respect to σ is:

(11) $\nabla_\sigma \varepsilon(u) = (A\ f_\sigma(u) - g)^T A\ diag(f'(u_i))_{i=\overline{1,n}}$.

The variation of u which leads to a constrained least-squares solution is on the descent direction of the gradient $\nabla_\sigma \varepsilon$:

(12) $\Delta u = -\eta \nabla_\sigma \varepsilon, \quad \eta > 0.$

The adaptive constrained 3D-reconstruction (AC3DR) algorithm starts with an arbitrary value for u, e.g. $u=0$, and iteratively modifies u with a quantity Δu computed as in the above equation.

The final regularised solution $f_\sigma(u)$ is thresholded and only the highest values are kept. This is the resulting object space reconstruction output by the current algorithm.

5. Hierarchical adaptive multiresolution strategy

The ill-conditioning of the problem is determined by the existence of very small singular values of the operator A, which virtually increase the dimension of the null-space of A in

numerical applications and make the solutions of the equation (4) be very unstable. The number of small singular values can be decreased by reducing the size of the matrix corresponding to A, so by considering only a low number of sampling points in the object space (lower than the number of pixels in the image). This operation has the drawback of inducing low resolution in the object reconstruction.

The conflict between ill-conditioning of the imaging operator and high resolution reconstruction requirement has been solved by adopting a ***hierarchical multiresolution strategy (HMS)*** based on zooming in the high interest regions. This approach consists of the following steps:

1) *Initial sampling and regularisation*: Sample a large domain $D^{(1)}$ using a course grid and obtain a (low) resolution reconstruction $\hat{I}_{D^{(1)}}$ from the equation $g = A_{D^{(1)}}\hat{I}_{D^{(1)}}$ using one of the regularisation schemes.
2) *Further sampling and regularisation:* Select the high interest regions $D_1,..., D_m$ of the object space from the initial reconstruction $\hat{I}_{D^{(1)}}$, and consider the rest of the space as determined. The following alternatives are possible:
   a) If the initial resolution is satisfactory and only the accuracy needs to be improved, perform further regularisation on $D^{(2)}=D_1\cup...\cup D_m$ using the restriction of the initial sampling to this sub-domain.
   b) If the resolution needs to be improved, sub-sample the union $D^{(2)}=D_1\cup...\cup D_m$ and obtain a finer resolution reconstruction $\hat{I}_{D^{(2)}}$ from the equation $g = A_{D^{(2)}}\hat{I}_{D^{(2)}}$.

If necessary, the procedure can be recursively applied until the object space is reconstructed at the desired resolution.

6. Simulation results

The hierarchical application of the projected Landweber method and of the adaptive constrained 3D-reconstruction algorithm has been tested on a set of computer generated images containing various object shapes. In the following, the results obtained at the reconstruction of an opaque cube are presented. For the simplicity of the presentation, planar sections parallel to the *Oyz*-plane of the object space reconstruction are shown. The object considered is perpendicular to the *Oyz*-plane, so its intersections with $x=x_0$ planes are similar. Hence, recovering one section means recovering the whole object. Figures 3-8 contain both the object reconstruction and the original object for comparison. The darker grey shades represent higher levels of intensity in these representations.

The initial PL method processing (HMS 1.)of the image of the visible sides of an opaque cube yielded the scene depicted in Figure 3. The object recovered at the initial sampling stage is symmetric and correctly positioned in space, and the two sides are clearly separated. The vicinity of the region where the cube was detected, marked by the dashed line in Figure 3, has been re-regularised (HMS 2.a)) using the restriction of the initial sampling grid to this zone and considering the rest of the space as determined. The result is shown in Figure 4 and provides a more accurately contoured object. Figure 5 depicts the second variant of refinement (HMS 2.b)) of the result shown in Figure 3. The rectangular

region enclosed in the dash-and-dotted line in Figure 3 has been resampled at increased resolution (i.e. zoomed in) and PL regularisation has been performed on it. The result, shown in Figure 5 is more accurate than the previous ones.

Figures 6-8 depict the corresponding processing stages performed using the AC3DR algorithm. The initial result of the AC3DR processing (HMS 1.) is depicted in Figure 6. Comparing it with the corresponding PL processing (Figure 3), it is noticed that the reconstruction displays higher contrast and better localisation. The localised AC3DR regularisation using the restriction of the initial sampling grid to the dashed region of Figure 6 (HMS 2.a)) provides the reconstruction shown in Figure 7. It follows accurately the contour of the object and is correctly positioned in space. Another processing, performed with increased resolution (HMS 2.b)) on the rectangular region marked by the dash-and-dotted line in Figure 6, is shown in Figure 8. The reconstruction is correctly shaped positioned, but the edges are less emphasised than in the previous reconstruction (Figure 7).

The simulations have shown that the two regularisation methods presented provide good reconstructions of the object space. A method of linking the edges in the reconstruction is necessary and it will be the object of further work.

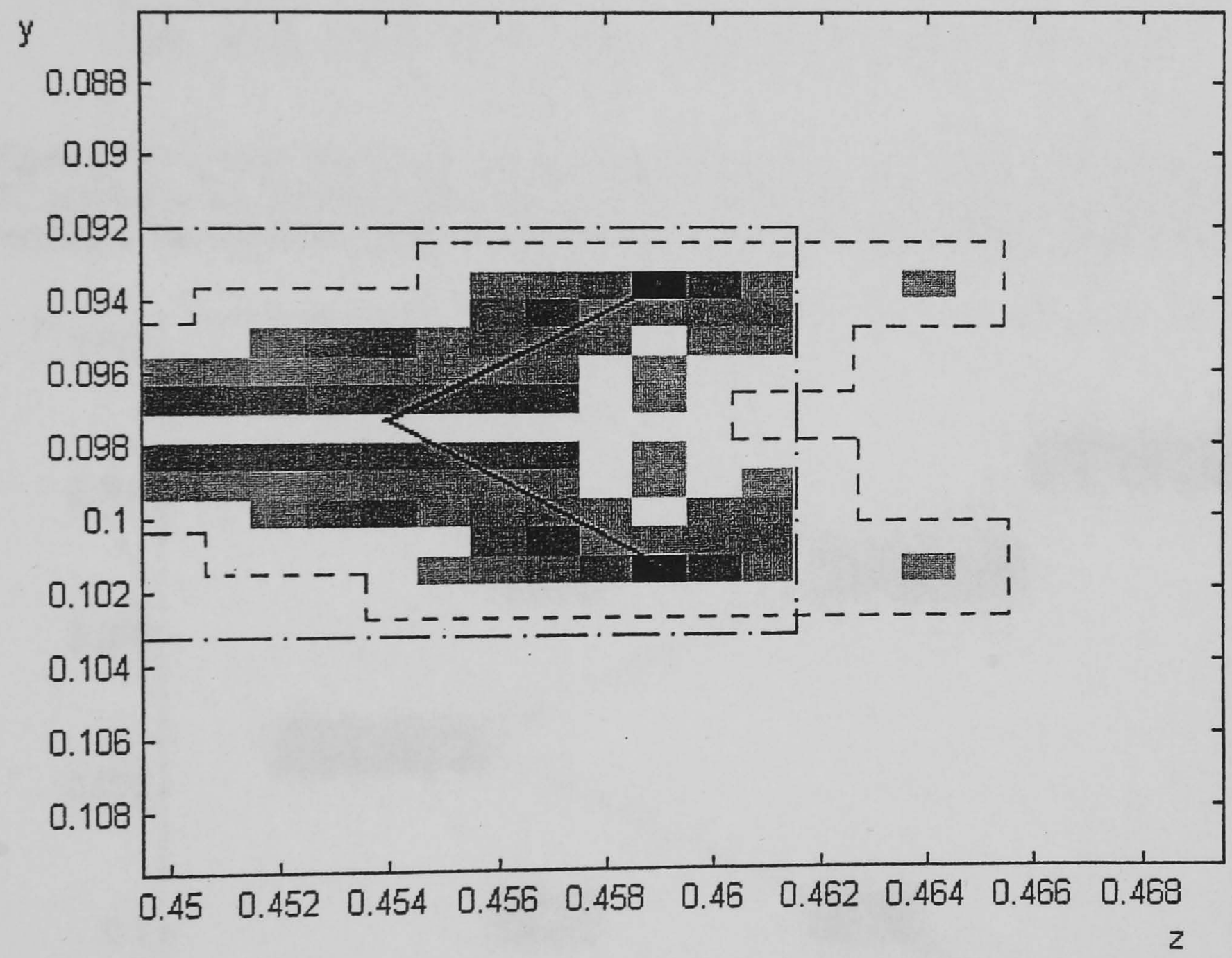


Figure 3. Initial regularisation of the object space obtained with the projected Landweber method. The actual contour of the object is shown in continuous line.

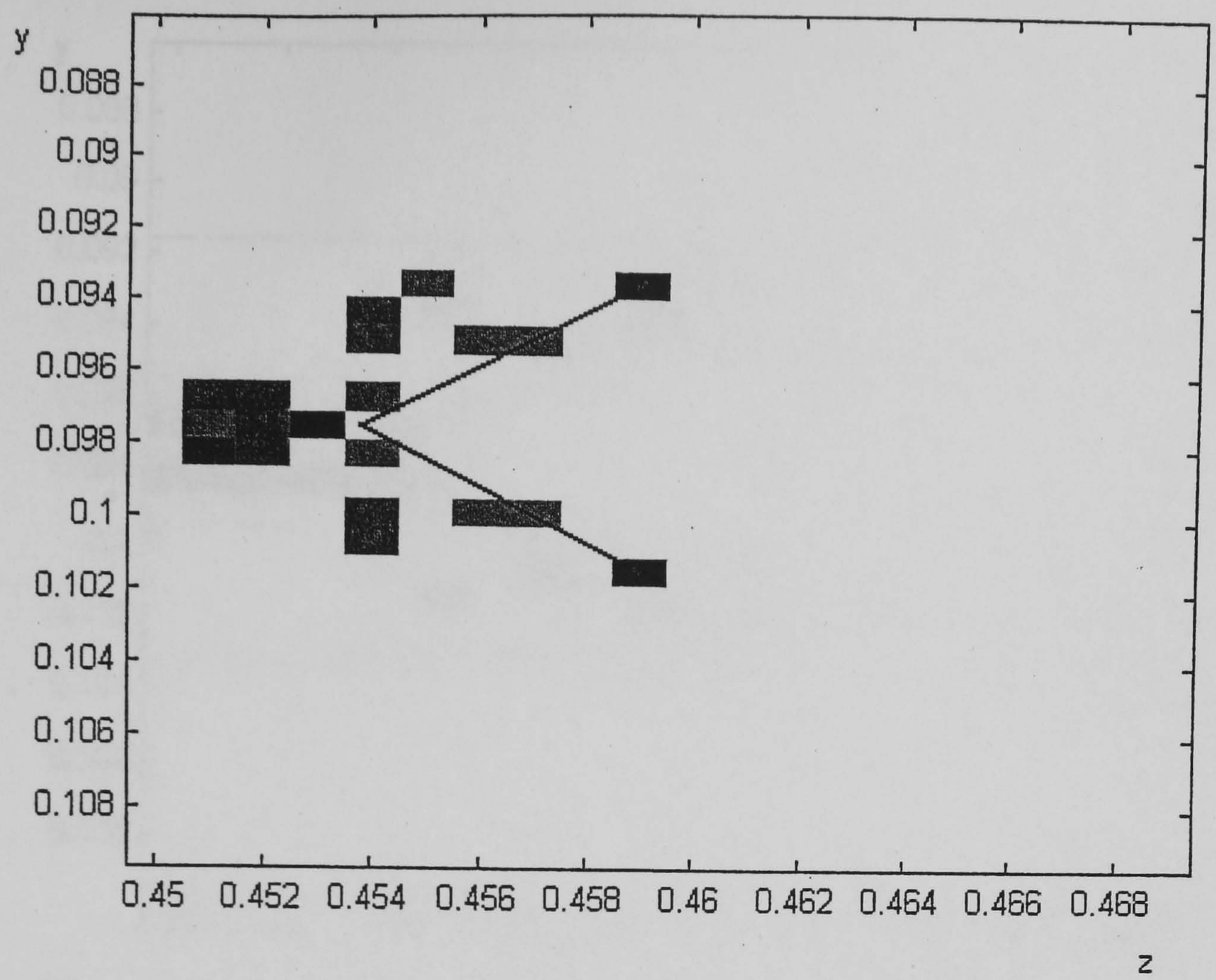


Figure 4. Transversal section through the reconstruction of the visible faces of an opaque cube using local PL regularisation around the high intensity region shown in Figure 3, preserving the initial sampling and resolution. The regularisation has been performed in the region enclosed by the dashed line in Figure 3.

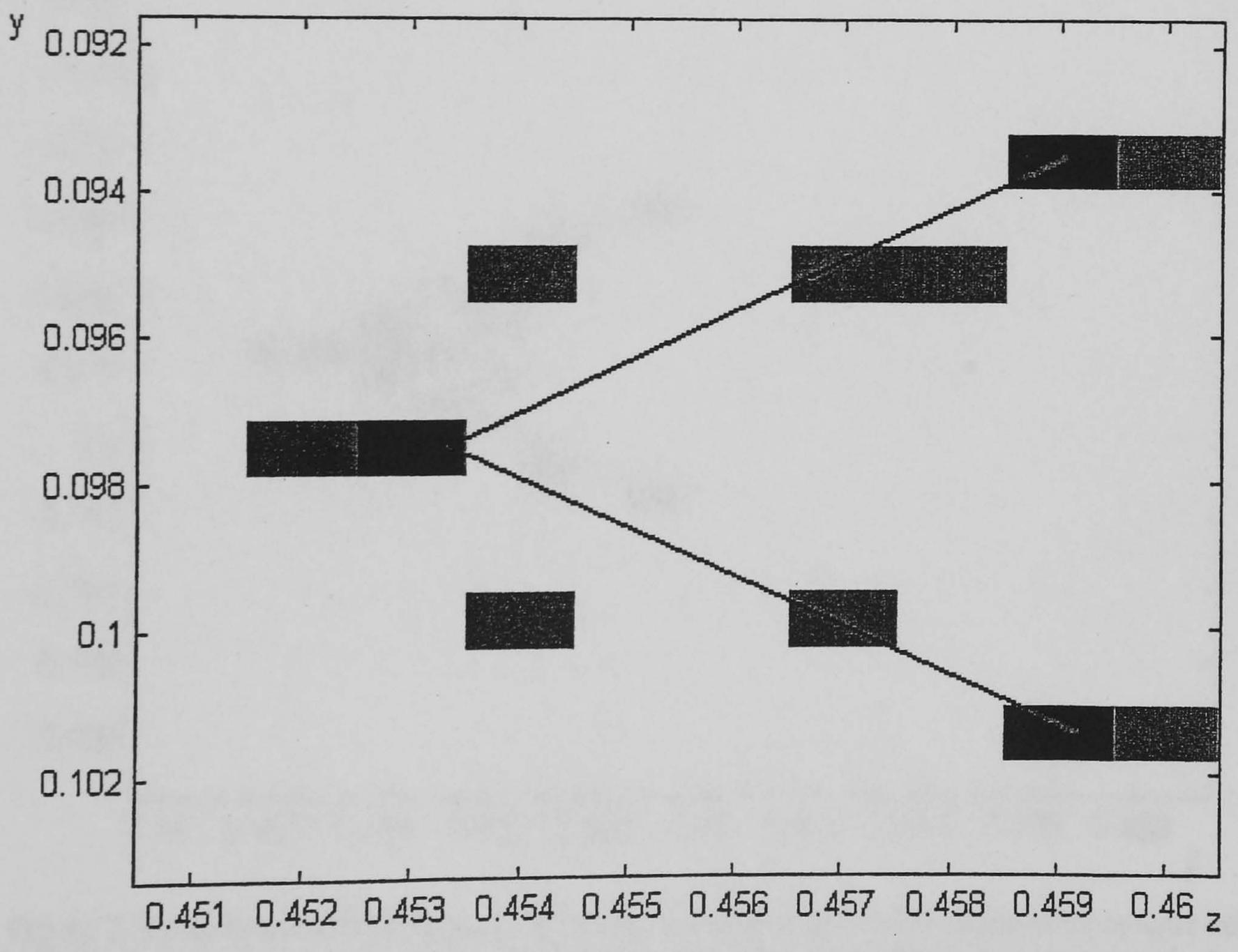


Figure 5. Transversal section through the reconstruction of the visible faces of an opaque cube using local PL regularisation around the high intensity region shown in Figure 3, using a finer sampling and obtaining higher resolution. The regularisation has been performed in the rectangular dash-and-dotted region in Figure 3.

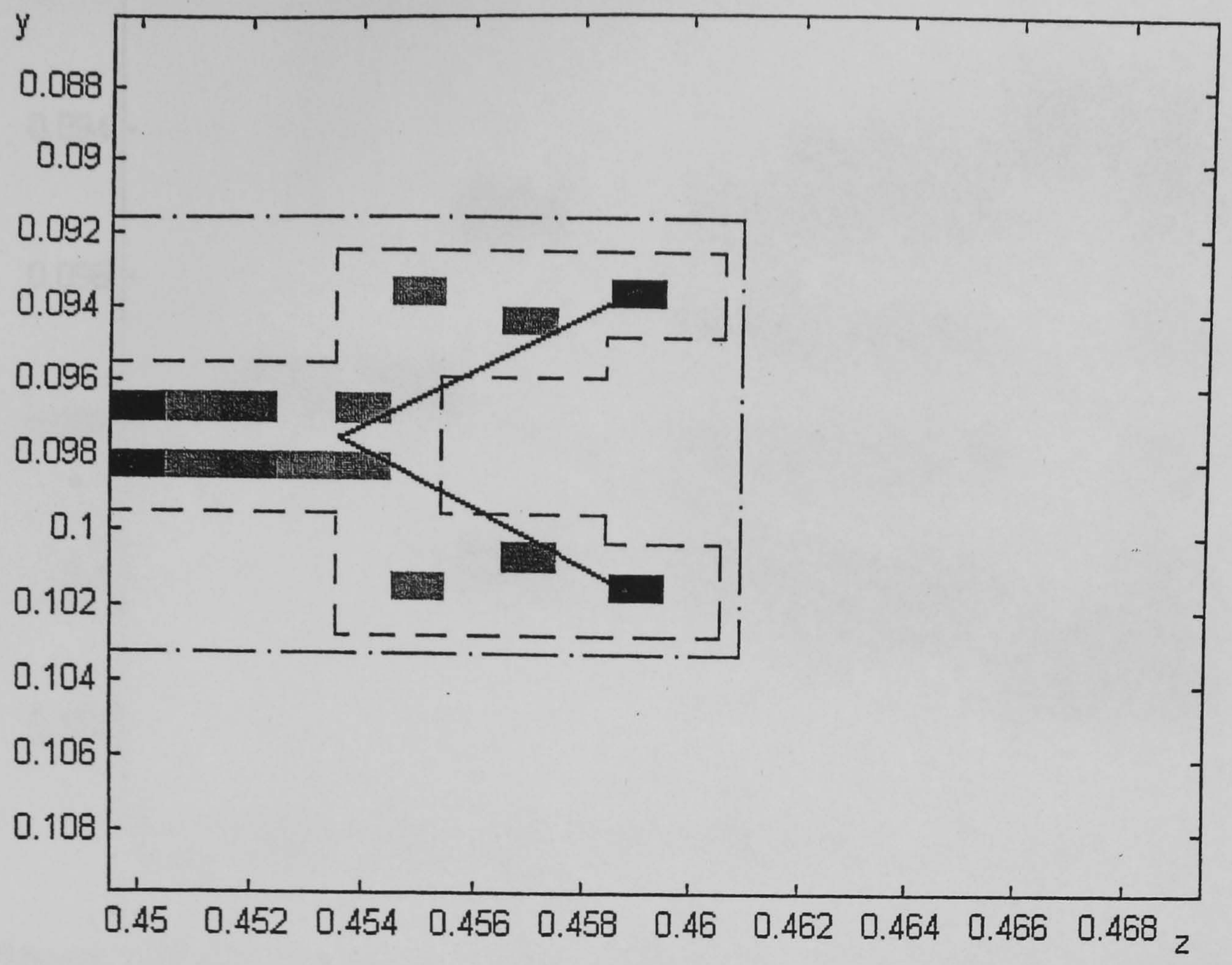


Figure 6. Initial regularisation of the object space obtained with the AC3DR algorithm. The actual contour of the object is shown in continuous line.

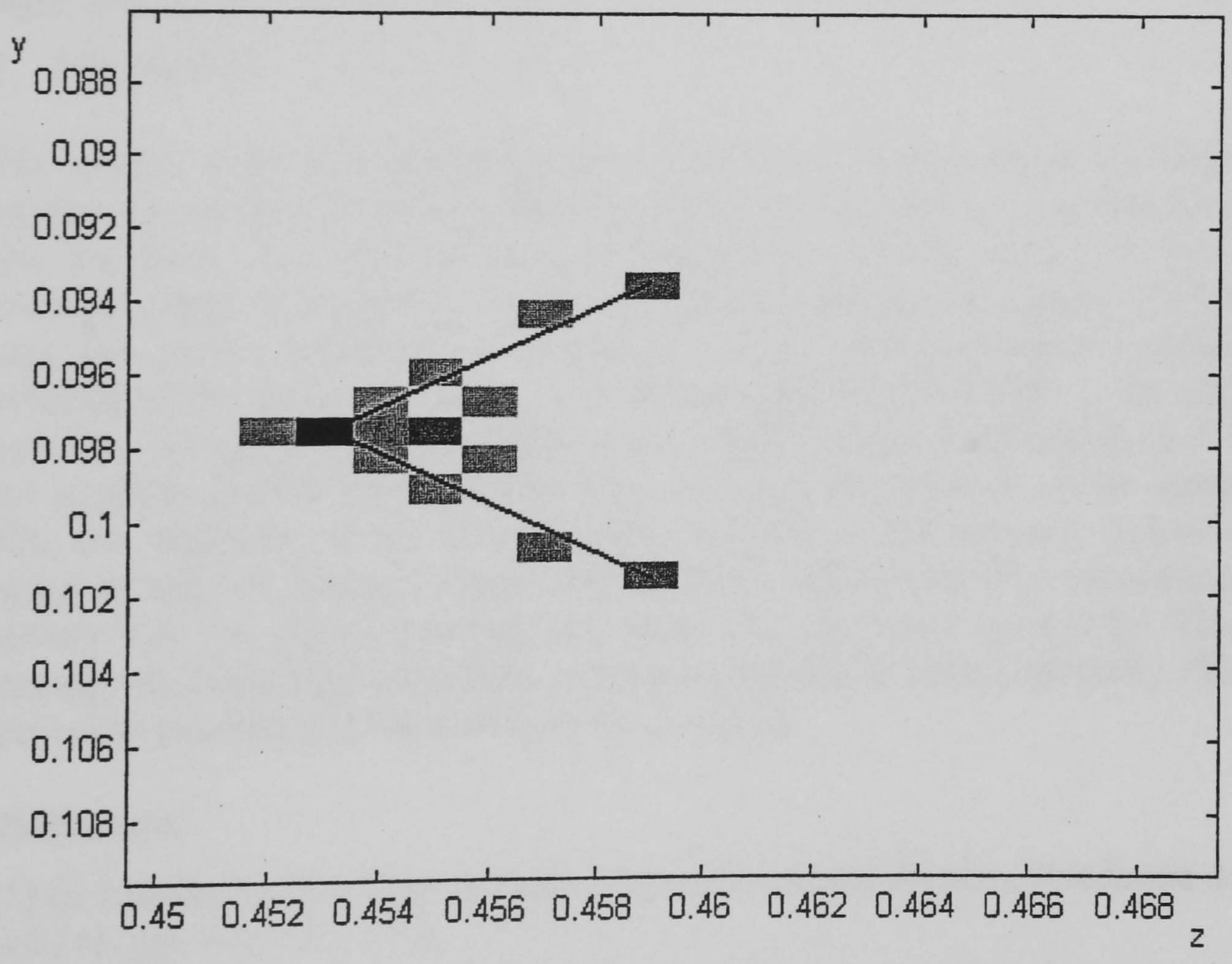


Figure 7. Transversal section through the reconstruction of the visible faces of an opaque cube using local AC3DR regularisation around the high intensity region shown in Figure 6, preserving the initial sampling and resolution. The regularisation has been performed in the region enclosed by the dashed line in Figure 6.

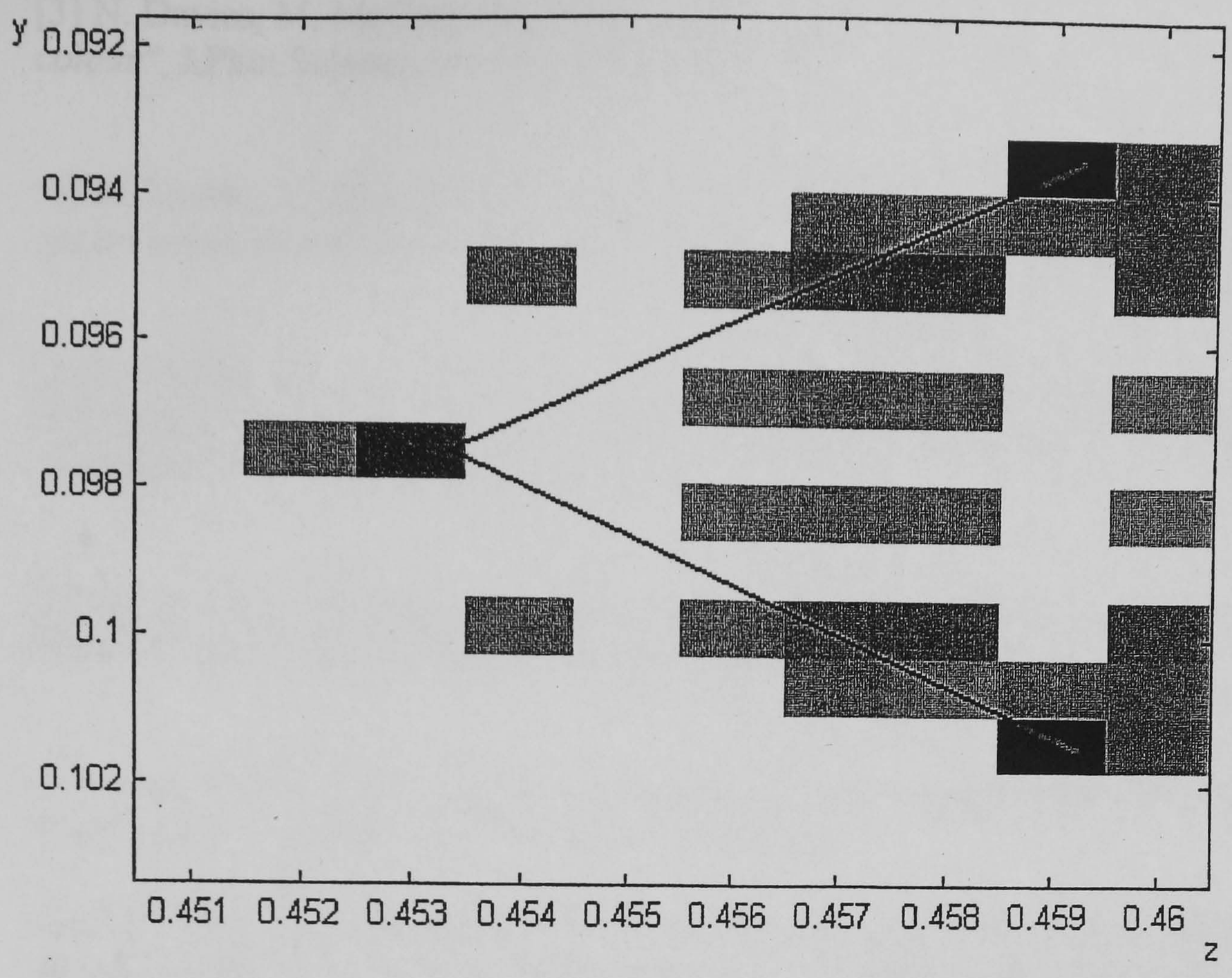


Figure 8. Transversal section through the reconstruction of the visible faces of an opaque cube using local AC3DR regularisation around the high intensity region shown in Figure 6, using a finer sampling and obtaining higher resolution. The regularisation has been performed in the rectangular dash-and-dotted region in Figure 6.

7. Conclusions

The present paper has approached the object reconstruction from 3D-integral images as an inverse problem, which proves ill-conditioned. In order to cure the ill-conditioning of the problem, the use of two regularisation methods able to provide realistic reconstructions is proposed. The first one is the projected Landweber (PL) method which uses the convex projection on the convex set of positive bounded functions to provide solution of the inverse problem. The second algorithm (AC3DR) is an adaptive scheme based on the use of sigmoid functions as a differentiable thresholding tool, which allows the gradient descent towards a positive least-squares solution of the imaging equation. The two methods can be hierarchically reapplied to the relevant regions of the object space in order to produce higher resolution reconstructions. The simulation results have shown that the spatial position and shape are recovered accurately. The hierarchical strategy produces high resolution reconstructions and is computationally efficient, as only the relevant regions of the scene are oversampled.

References

[1] G. Lippmann, "La Photographie integrale", Comtes Rendus, Academie des Sciences, vol.146, pp. 446-451, 1908.

[2] T. Okoshi, "Three dimensional imaging techniques", Academic Press, London, 1976.

[3] N. Davies, M. McCormick, *"Holoscopic imaging with true 3D-content in full natural colour"*, J.Phot.Science, vol. 40, pp. 46-49, 1992.

[4] N. Davies, M. McCormick, M. Brewin, *"Design and analysis of an image transfer system using microlens arrays"*, Opt. Eng., vol. 33-11, pp. 3624-3633, 1994.

[5] S. Manolache, A. Aggoun, M. McCormick, N. Davies, *"A mathematical model of a 3D-lenticular integral recording system"*, Proceedings of IEEE Vision, Modeling and Visualization, pp. 51-58, Erlangen, 1999.

[6] M. Bertero, P. Boccacci, *"Introduction to inverse problems in imaging"*, Institute of Physics Publishing, Bristol and Philadelphia, 1998.

[7] L. Landweber, *"An iteration formula for Fredholm integral equations of the first kind"*, Amer. J. Math., vol. 73, pp. 615-624, 1951.

[8] B. Eicke, *"Iteration methods for convexly constrained ill-posed problems in Hilbert space"*, Numer. Funct. Anal. and Optimiz., vol. 13-5&6, pp. 413-429, 1992.

[9] S. Manolache, M. McCormick, S.-Y. Kung, *"3D-object space reconstruction from planar recorded data of 3D-integral images"*, submitted for publication in Journal of VLSI Signal Processing Systems (Kluwer).